Decidability and complexity of the fragments of the modal logic of Allen's relations over the rationals

Decidability and Complexity of the Fragments of the Modal Logic of Allen's Relations over the Rationals[*]

D. Bresolin[a], D. Della Monica[b,c,d,e], A. Montanari[e], P. Sala[f], G. Sciavicco[g]

[a] *Dept. of Mathematics, University of Padova, Italy*
[b] *Istituto Nazionale di Alta Matematica "F. Severi" (INdAM), Italy*
[c] *Dept. of Electronic Engineering and Information Technologies, University of Napoli, Italy*
[d] *Sistemas Informaticos y Computación, Universidad Complutense de Madrid, Spain*
[e] *Dept. of Mathematics, Computer Science, and Physics, University of Udine, Italy*
[f] *Dept. of Computer Science, University of Verona, Italy*
[g] *Dept. of Mathematics and Computer Science, University of Ferrara, Italy*

Abstract

Interval temporal logics provide a natural framework for temporal reasoning about interval structures over linearly ordered domains, where intervals are taken as first-class citizens. Their expressive power and computational behaviour mainly depend on two parameters: the set of modalities they feature and the linear orders over which they are interpreted. In this paper, we consider all fragments of Halpern and Shoham's interval temporal logic HS with a decidable satisfiability problem over the rationals, and we provide a complete classification of them in terms of their expressiveness and computational complexity by solving the last few open problems.

Keywords: interval temporal logics, satisfiability, decidability, computational complexity

1. Introduction

Most temporal logics proposed in the literature assume a point-based structure of time. They have been successfully applied in a variety of fields,

ranging from the specification and verification of communication protocols to temporal data mining. However, a number of relevant application domains, such as, for instance, planning and synthesis of controllers and verification of multi-agent systems, would greatly benefit from the use of interval temporal logic, which has the ability to deal with advanced features that are neglected (accomplishments and temporal aggregations) or managed in an unsatisfactory way (durative actions and their temporal relationships) by point-based temporal logics.

Temporal logics with interval-based semantics have been proposed as suitable formalisms for the specification and verification of hardware [2], software [3], real-time [4], and multi-agent systems [5]. Moreover, successful implementations of interval-based systems can be found in the areas of machine learning, e.g., the adaptive learning system TERENCE [6], that provides a support to poor comprehenders and their educators, is based on Allen's interval algebra [7], and real-time data systems, e.g., the algorithm RISMA [8], for performance and behaviour analysis of real-time data systems, is based on Halpern and Shoham's modal logic of Allen's relations [9].

The variety of binary relations between intervals in a linear order was first studied by Allen [7], who investigated their use in systems for time management and planning; Allen identified 12 different relations between two intervals in a linear order, often called *Allen's relations* [7]: the six relations R_A (*meets* or *adjacent*), R_L (*after* or *later*), R_B (*starts* or *begins*), R_E (*finishes* or *ends*), R_D (*during*), and R_O (*overlaps*), depicted in Fig. 1, and their inverses, that is, $R_{\overline{X}} = (R_X)^{-1}$, for each $X \in \{A, L, B, E, D, O\}$. In [9], Halpern and Shoham introduced and studied the (full) modal logic of Allen's relations (HS for short), that has one modality for each Allen relation, showing that the satisfiability problem for HS is highly undecidable over most (classes of) linear orders. This result motivated the search for syntactic fragments of HS offering a good balance between expressiveness and computational complexity. During the last decade, a systematic analysis of HS fragments has been carried out to characterize the complexity of their satisfiability problem [10, 11] as well as their relative expressive power [11, 12, 13]. Such an analysis clearly showed that expressiveness and complexity (decidability) of HS fragments depend on both the modalities they feature and the class of linear orders over which they are interpreted. The case of the logic of the sub-interval relation (Allen relation *during*) is paradigmatic in this respect: it is decidable (PSPACE-complete) over the rationals and

HS modalities	Allen's relations	Graphical representation
$\langle A \rangle$	$[x,y]R_A[x',y'] \Leftrightarrow y = x'$	
$\langle L \rangle$	$[x,y]R_L[x',y'] \Leftrightarrow y < x'$	
$\langle B \rangle$	$[x,y]R_B[x',y'] \Leftrightarrow x = x', y' < y$	
$\langle E \rangle$	$[x,y]R_E[x',y'] \Leftrightarrow y = y', x < x'$	
$\langle D \rangle$	$[x,y]R_D[x',y'] \Leftrightarrow x < x', y' < y$	
$\langle O \rangle$	$[x,y]R_O[x',y'] \Leftrightarrow x < x' < y < y'$	

Figure 1: Allen's interval relations and the corresponding HS modalities.

undecidable over finite/discrete linear orders.

This paper aims at providing a complete classification of decidable HS fragments, interpreted over the linear order of the rational numbers $\mathbb{Q}$ (equivalently, the class of all dense linear orders), with respect to both their expressiveness and the complexity of their satisfiability problem. An analogous classification has been provided in [11] for the class of finite linear orders, the class of discrete linear orders, the linear order of the natural numbers $\mathbb{N}$, and the linear order of the integers $\mathbb{Z}$. As for the class of all linear orders, a complete classification of HS fragments with respect to their expressiveness can be found in [12], while decidability and/or complexity of the satisfiability problem are still unknown for some meaningful fragments, the most interesting one being probably the logic of Allen relation *during*.

Related work. While a complete classification of the expressiveness of HS fragments over $\mathbb{Q}$ was provided in [12], some fundamental tiles of their classification with respect to the computational complexity of the satisfiability problem are still missing. The aim of this paper is to complete the picture by filling all the gaps.

Known results about decidability and complexity of the satisfiability problem for HS fragments over $\mathbb{Q}$ can be summarized as follows.

The set of undecidable fragments of HS consists of *(i)* HS fragments containing, possibly as definable, modality $\langle O \rangle$ (Allen relation *overlaps*) or $\langle \overline{O} \rangle$ (*overlapped by*) [10], *(ii)* HS fragments containing, possibly as definable, modality $\langle B \rangle$ (*begins*) or $\langle \overline{B} \rangle$ (*begun by*) together with modality $\langle E \rangle$ (*ends*) or $\langle \overline{E} \rangle$ (*ended by*) [14], and *(iii)* HS fragments containing, possibly as definable, modality $\langle A \rangle$ (*meets*) or $\langle \overline{A} \rangle$ (*met by*) together with modality $\langle D \rangle$

3

($during$) or $\langle \overline{D} \rangle$ ($contains$) [10].

The set of decidable fragments of HS contains the logics $\mathsf{A\overline{A}B\overline{B}L\overline{L}}$ (featuring Allen's relations *meets*, *met by*, *begins*, *begun by*, *before*, and *later*), $\mathsf{A\overline{A}E\overline{E}L\overline{L}}$ (*meets*, *met by*, *ends*, *ended by*, *before*, and *later*), $\mathsf{B\overline{B}D\overline{D}L\overline{L}}$ (*begins*, *begun by*, *during*, *contains*, *before*, and *later*), and $\mathsf{E\overline{E}D\overline{D}L\overline{L}}$ (*ends*, *ended by*, *during*, *contains*, *before*, and *later*), and their fragments.[1] The fragment $\mathsf{A\overline{A}B\overline{B}}$ has been shown to be non-primitive recursive in [15] (the same holds for the fragment $\mathsf{A\overline{A}E\overline{E}}$ by symmetry). Known complexity results for its fragments (resp., for the fragments of $\mathsf{A\overline{A}E\overline{E}}$) are as follows: $\mathsf{AB\overline{B}L}$ (resp., $\mathsf{\overline{A}E\overline{E}L}$) is ExpSpace-complete [16], while A, $\mathsf{\overline{A}}$, and $\mathsf{A\overline{A}}$ are NExpTime-complete [17]. The PSpace-completeness of the logic D of the sub-interval relation, and of some variants of it, has been proved in [18, 19]. Such a result has been improved in [20], where the authors showed that PSpace-completeness can actually be lifted to the logic $\mathsf{B\overline{B}D\overline{D}L\overline{L}}$ (and to the symmetric fragment $\mathsf{E\overline{E}D\overline{D}L\overline{L}}$).

Main contributions. In this paper, we focus our attention on the set of decidable fragments of HS and provide a fine-grained characterization of their computational complexity that answers all the remaining open issues. First, we show that the satisfiability problem for the HS fragment $\mathsf{B\overline{B}L\overline{L}}$, which is obtained from $\mathsf{B\overline{B}D\overline{D}L\overline{L}}$ by removing the modalities for Allen's relations *during* and *contains*, is NP-complete. Such a result immediately propagates to all the fragments of $\mathsf{B\overline{B}L\overline{L}}$. Then, we prove that the satisfiability problem for all HS fragments containing either the pair of modalities $\langle A \rangle$ (*meets*) and $\langle B \rangle$ (*begins*) or the pair $\langle A \rangle$ (*meets*) and $\langle \overline{B} \rangle$ (*begun by*) is ExpSpace-hard. Finally, we show that the satisfiability problem for all HS fragments containing either the pair of modalities $\langle \overline{A} \rangle$ (*met by*) and $\langle B \rangle$ (*begins*) or the pair $\langle \overline{A} \rangle$ (*met by*) and $\langle \overline{B} \rangle$ (*begun by*) is not primitive recursive (we provide a non-primitive recursive lower bound).

As we already pointed out, the present work is a considerably revised and extended version of [1]. Its main novelties with respect to [1] can be summarized as follows.

1. We work out in all the details the proofs of the NP-completeness of $\mathsf{B\overline{B}L\overline{L}}$ (Section 5) and of the non-primitive recursive lower bound for

---

[1]In fact, modalities $\langle L \rangle$ (Allen relation *later*) and $\langle \overline{L} \rangle$ (*before*) can be easily defined in terms of modalities $\langle A \rangle$ and $\langle \overline{A} \rangle$, respectively, and thus $\mathsf{A\overline{A}B\overline{B}L\overline{L}}$ (resp., $\mathsf{A\overline{A}E\overline{E}L\overline{L}}$) is as expressive as $\mathsf{A\overline{A}B\overline{B}}$ (resp., $\mathsf{A\overline{A}E\overline{E}}$).

$\overline{\text{AB}}$ and $\overline{\text{A}}\overline{\text{B}}$ (Section 7). Both results are stated in [1], where only a very high-level account of the main proof ideas is given (as a matter of fact, we changed a lot the structure of the proof of the first result).

2. We prove the EXPSPACE-hardness of both AB and $\text{A}\overline{\text{B}}$ (Section 6). Such a result was stated in [1], but no proof was given.

3. In order to show the potentialities of the considered logics, we use them to model and to reason about a concrete application domain, namely, the Chicago's bike sharing program [21] (Section 3).

4. We discuss the relationships among different (classes of) dense temporal domains (the linear order of the rationals, the linear order of the reals, the class of all dense linear orders) with respect to the satisfiability problem for HS fragments. As it turns out, HS fragments behave on $\mathbb{Q}$ exactly as they do on the class of all dense linear orders, but, somehow surprisingly, differences may occur when moving from $\mathbb{Q}$ to $\mathbb{R}$.

Structure of the paper. In Section 2, we introduce syntax and semantics of HS and its fragments. Then, in Section 3 we show how HS fragments can be successfully exploited to reason about a concrete application domain, namely, the Chicago's bike sharing program. Next, in Section 4, putting together known and new results (that will be proved in the subsequent sections), we give the complete picture of the computational complexity of the satisfiability problem for the decidable HS fragments interpreted over $\mathbb{Q}$. Section 5, Section 6, and Section 7 respectively focus on the subclasses of NP-complete, EXPSPACE-complete, and non-primitive recursive fragments of HS. Finally, in Section 8, we briefly discuss the relationships among the satisfiability problem for HS fragments interpreted over $\mathbb{Q}$, $\mathbb{R}$, and the class of (all) dense linear orders. Conclusions provide an assessment of the work done and briefly discuss related and future work.

2. Syntax and semantics of HS and its fragments

Let $<$ be the usual relation of total order over the set $\mathbb{Q}$ of the rational numbers. An *interval* over $\mathbb{Q}$ is an ordered pair $[x, y]$, where x, y are rational numbers and $x \leq y$, that is, either $x = y$ or $x < y$. An interval is called a *point interval* if $x = y$ and a *strict interval* if $x < y$. In this paper, we assume the *strict* semantics, that is, we exclude point intervals and only consider strict intervals. The adoption of the strict semantics, excluding point intervals, instead of the *non-strict* one, which includes them, conforms to the definition

of interval given by Allen in [7], but differs from the one given by Halpern and Shoham in [9]. It has at least two strong motivations: first, a number of representation paradoxes arise when the non-strict semantics is adopted, due to the presence of point intervals, as pointed out in [7]; second, when point intervals are included there seems to be no intuitive semantics for interval relations that makes them both pairwise disjoint and jointly exhaustive.

If we exclude the identity relation, there are 12 different relations between two strict intervals in a linear order, often called *Allen's relations* [7]: the six relations R_A (*meets* or *adjacent*), R_L (*after* or *later*), R_B (*starts* or *begins*), R_E (*finishes* or *ends*), R_D (*during*), and R_O (*overlaps*), depicted in Fig. 1, and their inverses, that is, $R_{\overline{X}} = (R_X)^{-1}$, for each $X \in \{A, L, B, E, D, O\}$.

We interpret interval structures as Kripke structures with Allen's relations playing the role of the accessibility relations. Thus, we associate a modality $\langle X \rangle$ with each Allen relation R_X. For each $X \in \{A, L, B, E, D, O\}$, the *transpose* of modality $\langle X \rangle$ is modality $\langle \overline{X} \rangle$, corresponding to the inverse relation $R_{\overline{X}}$ of R_X. Halpern and Shoham's logic HS [9] is a multi-modal logic with formulae built from a finite, non-empty set $\mathcal{AP}$ of atomic propositions (also referred to as proposition letters), the propositional connectives $\vee$ and $\neg$, and a modality for each Allen relation. With every subset $\{R_{X_1}, \ldots, R_{X_k}\}$ of these relations, we associate the fragment $\mathsf{X_1 X_2 \ldots X_k}$ of HS, whose formulae are defined by the grammar:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \langle X_1 \rangle \varphi \mid \ldots \mid \langle X_k \rangle \varphi,$$

where $p \in \mathcal{AP}$. The other propositional connectives and constants, e.g., $\wedge$, $\rightarrow$, and $\top$, as well as the dual modalities, e.g., $[A]\varphi \equiv \neg\langle A \rangle \neg\varphi$, can be derived in the standard way.

The (strict) semantics of HS is given in terms of *interval models* $M = \langle \mathbb{I}(\mathbb{Q}), V \rangle$, where $\mathbb{I}(\mathbb{Q})$ is the set of all (strict) intervals over $\mathbb{Q}$ and V is a *valuation function* $V : \mathcal{AP} \rightarrow 2^{\mathbb{I}(\mathbb{Q})}$, which assigns to each atomic proposition $p \in \mathcal{AP}$ the set of intervals $V(p)$ on which p holds. The *truth* of a formula on a given interval $[x, y]$ in an interval model M is defined by structural induction on formulae as follows:

- $M, [x, y] \Vdash p$ if and only if $[x, y] \in V(p)$, for each $p \in \mathcal{AP}$;
- $M, [x, y] \Vdash \neg\psi$ if and only if it is not the case that $M, [x, y] \Vdash \psi$;
- $M, [x, y] \Vdash \varphi \vee \psi$ if and only if $M, [x, y] \Vdash \varphi$ or $M, [x, y] \Vdash \psi$;
- $M, [x, y] \Vdash \langle X \rangle \psi$ if and only if there exists an interval $[x', y']$ such that $[x, y] R_X [x', y']$ and $M, [x', y'] \Vdash \psi$, for each modality $\langle X \rangle$.

Given an interval model M and a formula φ, we say that M *satisfies* φ if there is an interval $[x, y]$ in $\mathbb{I}(\mathbb{Q})$ such that $M, [x, y] \Vdash \varphi$. We say that φ is *satisfiable* if there exists an interval model that satisfies it.[2]

We sometimes omit the model when it is clear from the context, and thus we write, e.g., $[x, y] \Vdash \varphi$ instead of $M, [x, y] \Vdash \varphi$. In the analysis of the various fragments of HS, we shall often make use of a *global* modality, denoted by $[G]$, which can be defined in terms of basic HS modalities. While the intended semantics of $[G]$ is the same across the various considered fragments, namely, it forces a given property to hold over all intervals belonging to a relevant set (the definition of such a set will be made formal in due course), its formal definition and the one of relevant set of intervals depend on the modalities featured by the fragment under consideration.

Formulae of HS are interpreted over a class of interval models. We identify the class of interval models over linear orders in $\mathcal{C}$ with the class $\mathcal{C}$ itself. Thus, we shall use, for example, the expression 'formulae of HS are interpreted in the class $\mathcal{C}$ of linear orders' instead of the extended one 'formulae of HS are interpreted over the class of interval models over linear orders in $\mathcal{C}$'. Among others, we mention the following relevant (classes of) linear orders: *(i)* the class of *all* linear orders Lin; *(ii)* the class of all *dense* linear orders Den, i.e., those in which for every pair of different points there exists at least one point in between them; *(iii)* the class of all *weakly discrete* linear orders WDis, i.e., those in which every element, apart from the greatest one, if it exists, has an immediate successor, and every element, other than the least one, if it exists, has an immediate predecessor; *(iv)* the class of all *strongly discrete* linear orders Dis, i.e., those in which for every pair of different points there are only finitely many points in between them; *(v)* the class of all *finite* linear orders Fin, i.e., those having only finitely many points; *(vi)* the singleton classes consisting of the standard linear orders over $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}$, and $\mathbb{N}$.

The *mirror image* (or, simply, *mirror*) of a fragment $\mathcal{F}$ is obtained by simultaneously substituting $\langle \overline{A} \rangle$ for $\langle A \rangle$, $\langle E \rangle$ for $\langle B \rangle$, $\langle \overline{E} \rangle$ for $\langle \overline{B} \rangle$, $\langle \overline{O} \rangle$ for

---

[2]It is worth pointing out that the distinctive feature of a really interval-based temporal logic is that its formulae express properties of pairs of time points rather than of single time points and are evaluated as sets of such pairs, that is, as binary relations. As a consequence, apart from some very special cases, e.g., the logic of Allen's relations *begins* and *begun by*, there is no a reduction of their satisfiability/validity problem to the satisfiability/validity problem for (fragments of) monadic second-order logic. This means, in particular, that Rabin's theorem cannot be exploited in the context of interval temporal logics.

$\langle O \rangle$, $\langle \overline{L} \rangle$ for $\langle L \rangle$, and the other way around (the mirror image of a formula is defined analogously). When interpreted over left/right symmetric classes of structures, that is, classes $\mathcal{C}$ such that if $\mathcal{C}$ contains a linear order $\mathbb{D} = \langle D, \prec \rangle$, then it also contains a linear order isomorphic to its dual linear order $\mathbb{D}^d = \langle D, \succ \rangle$, where $\succ$ is the inverse of $\prec$, such as the class Den, all computational properties of a fragment are preserved by its mirror image. Thanks to such a property, for each pair of mirror fragments, we can safely restrict our attention to only one of them. In the following, we focus on the linear order of the rational numbers $\mathbb{Q}$ (equivalently, the class Den of all dense linear orders).

3. Interval temporal logics at work

In this section, we illustrate the potentialities of the various HS fragments studied in this paper by means of a concrete application example taken from the domain of temporal databases [22].

Generally speaking, in comparison to standard point-based linear temporal logic, interval temporal logic has two advantages and one drawback. First of all, interval temporal logic allows one to express many relevant properties in a more natural and compact way than point-based one. The following simple example, due to Jakub Michaliszyn, gives a nice exemplification of the claim. Consider the property: "there is always a banquet during a conference". In interval temporal logic (HS fragment D), it is captured by the formula: $conference \rightarrow \langle D \rangle \ banquet$. The same property can be expressed in Linear Temporal Logic (LTL), interpreted in a discrete temporal domain, by the formula:

$$conference_{start} \longrightarrow$$
$$X(\neg conference_{end}U(banquet_{start} \wedge (\neg conference_{end}U(banquet_{end} \wedge$$
$$\neg conference_{end} \wedge F \ conference_{ends}))))$$

Second, as we have already pointed out in the introduction, there are temporal conditions which are inherently interval-based, and thus cannot be properly dealt with in point-based temporal logic. This is the case with telic statements [23], that is, statements that express goals or accomplishments, like the statement: "The airplane flew from Venice to Toronto" (an encoding of such a statement in the HS fragment $AB\overline{B}$ can be found in [16, Sect. II.B]). Similarly, interval temporal logic makes it possible to express statements like:

"the average speed of the mobile device in the considered section must not exceed the specified threshold", which involves a temporal aggregation and thus cannot be associated with the single time points.[3] The disadvantage of interval temporal logic is the computational complexity of its satisfiability problem. Luckily, there are various computationally well-behaved fragments of HS, which are expressive enough to express the requirements of realistic application domains, as shown by the considered case study.

The application domain. The example is inspired by Divvy, a system for the management of bike sharing developed by the city of Chicago. The system consists of a fleet of bikes that are locked into a network of docking stations located throughout the city. Bikes can be rented from and returned to any station in the city, creating a network of trips with many possible combinations of starting and ending points. Anonymized trip data are stored in a temporal database and openly available through the "Divvy Data Challenge" program [21]. Here we consider the slightly simplified version of the Data Challenge trip database depicted in Table 1. The database consists of two tables: a table *Trips*, maintaining information about user rentals, and a table *Maintenance*, storing data about repairs and other maintenance activities on bikes. We assume that repairs and maintenance are carried out at a special station (with identifier 1) and that bikes are regularly collected from the stations in the city and brought to the repair station. At the end of the repair activity, the bike can be rented again by the customers from station 1 (it is not brought back to the station from which it was collected). If a bike is collected for repair at station 1, no data is stored in the table *Maintenance* (we record only maintenance trips originating from a non-repair station). Every trip (resp., repair) is represented by the following data:

- a unique *trip_id* (resp., *repair_id*);

---

[3]A systematic comparison of the expressive power of interval temporal logics and point-based ones, over discrete temporal domains, has been done in [24], where it is shown that under the *homogeneity assumption* [25], according to which a proposition letter holds over an interval if and only if it holds at all its points (notice that such an assumption considerably weakens the expressive power of an interval temporal logic as it prevents it from modeling temporal features like accomplishments and temporal aggregations that do not satisfy the assumption), LTL and HS have the same expressive power, but the latter is provably at least exponentially more succinct than the former (in fact, the HS fragment AB suffices to capture LTL). As a matter of fact, the comparison has been done in the model checking setting, but it can be easily adapted to the satisfiability checking one.

Trips					
trip_id	starttime	stoptime	bikeid	from_station_id	to_station_id
4118	2013-06-27 12:11	2013-06-27 12:16	316	85	28
4275	2013-06-27 14:44	2013-06-27 14:45	64	32	32
4291	2013-06-27 14:58	2013-06-27 15:05	433	32	19
4316	2013-06-27 15:06	2013-06-27 15:09	123	19	19
4342	2013-06-27 15:13	2013-06-27 15:27	852	19	55
4480	2013-06-27 19:40	2013-06-27 22:28	27	340	46

Maintenance					
repair_id	starttime	stoptime	bikeid	from_station_id	to_station_id
5327	2013-06-28 09:05	2013-06-28 10:15	594	27	1
5335	2013-06-28 09:14	2013-06-28 10:41	227	26	1
5346	2013-06-28 09:26	2013-06-28 14:25	118	74	1
5353	2013-06-28 09:35	2013-06-28 09:50	226	24	1

Table 1: The bike sharing system database.

- an interval [*starttime*, *stoptime*], that keeps track of the time period during which the trip (resp., repair) takes place;
- a *bike_id*, that associates a bike with a trip (resp., repair);
- a pair of station identifiers (*from_station_id*, *to_station_id*), that respectively designate the station from which the bike is collected and the station to which it is returned.

The formalization. Interval temporal logics can be used to reason about database knowledge in at least two different ways: as *query languages*, to extract information from a given database instance, and as *specification languages*, to express functional dependencies, integrity constraints, and other requirements that all database instances must satisfy. Here we focus on the latter, showing how different HS fragments can be successfully exploited as specification languages. In general, once the requirements of a given (database) application domain have been collected and formalized by domain experts, the first problem the designer must deal with is *consistency*, that is, the problem of checking whether or not these requirements can actually be met by a concrete instance of the database. In its most general formulation, the consistency problem can be expressed as the problem of checking whether the set of formulae representing the requirements is satisfiable or not.

In the following, we assume the domain of the *starttime* and *stoptime* timestamps to be the set of nonnegative rationals $\mathbb{Q}^{\geq 0}$ and the sets of bike

and station identifiers to be finite, and we ignore trip and repair identifiers. To express the requirements, we make use of the following proposition letters:

- proposition letters of the form $trip(i, j, k)$, representing a trip of bike i from station j to station k;
- proposition letters of the form $repair(i, j, k)$, representing a bike repair (bike i was collected at station j and brought to station k to be repaired).

Every instance of the database corresponds to an HS model, where an interval $[x, y]$ satisfies $trip(i, j, k)$ if and only if there is a tuple in the table *Trips* with $starttime = x$, $stoptime = y$, $bikeid = i$, $from\_station\_id = j$, and $to\_station\_id = k$, and it satisfies $repair(i, j, k)$ if and only if there is a tuple in the table *Maintenance* with $starttime = x$, $stoptime = y$, $bikeid = i$, $from\_station\_id = j$, and $to\_station\_id = k$.

To simplify formulae, we shall use $trip(i)$ (resp., $repair(i)$) as a shorthand for $\bigvee_{j,k} trip(i, j, k)$ (resp., $\bigvee_{j,k} repair(i, j, k)$). Similarly, we shall write $trip(i, j)$ (resp., $repair(i, j)$) for $\bigvee_k trip(i, j, k)$ (resp., $\bigvee_k repair(i, j, k)$).

Finally, for the rest of the section, let $[G]$ (*global modality*) be defined as:

$$[G]\varphi = \varphi \wedge [B]\varphi \wedge [\overline{B}]\varphi \wedge [\overline{B}][D]\varphi,$$

which can be interpreted as follows: for each formula φ, $[G]\varphi$ is true over an interval $[x, y]$ if, and only if, φ is true over all intervals $[x', y']$, with $x' \geq x$.

Trip and repair requirements. We split the set of requirements into three subsets: a set of requirements on bike trips, a set of requirements on bike repairs, and a set of requirements that involve both of them. In this paragraph, we show how to encode them in suitable fragments of HS; in the next one, we shall discuss the problem of checking their satisfiability (consistency).

Since bike trips and repairs must satisfy, among others, some very similar requirements, we preliminarily introduce some general parametric formulae and then suitably instantiate them in the two cases. Let φ be an arbitrary propositional formula, that is, a formula devoid of temporal modalities. The following formula constrains intervals over which φ holds to be disjoint:

$$disj(\varphi) \;=\; [G]\big(\varphi \rightarrow (\neg\langle B\rangle\varphi \wedge \neg\langle D\rangle\varphi \wedge \neg\langle E\rangle\varphi \wedge \neg\langle O\rangle\varphi \wedge \neg\langle A\rangle\varphi)\big) \quad (1)$$

The next formula forces a non-infinitesimal delay between consecutive intervals over which φ holds:

$$del(\varphi) \;=\; [G]\big(\varphi \rightarrow \langle A\rangle(\neg\langle D\rangle\varphi \wedge \neg\langle E\rangle\varphi \wedge \neg\langle O\rangle\varphi)\big) \quad (2)$$

11

Such a condition makes it evident the advantages of the choice of a dense temporal domain over a discrete one, as it relieves us from the necessity to introduce an artificial discretization of the temporal domain.

The first requirement we enforce on table *Trip* is that at any time point a given bike can belong to at most one trip and that there is a delay between a trip and the next one. Formally, for each bike identifier i, (a) every interval satisfies at most one $trip(i, j, k)$, (b) pairs of intervals respectively satisfying $trip(i, j, k)$ and $trip(i, l, m)$ must be disjoint, and (c) there must be a non-infinitesimal delay between the end of a trip and the beginning of the next one. In full HS, the above requirement can be expressed by means of the following pair of formulae:

$$\bigwedge_{i,j,k} [G]\big(trip(i, j, k) \to \neg \bigvee_{(j,k) \neq (l,m)} trip(i, l, m)\big) \tag{3}$$

$$\bigwedge_i disj(trip(i)) \wedge \bigwedge_i del(trip(i)) \tag{4}$$

Arguably, full HS makes it possible to express this requirement in a pretty easy and natural fashion. Unfortunately, formula (4) belongs to the HS fragment $\mathsf{AB\overline{B}DEO}$ (modality $\langle\overline{B}\rangle$ occurs in the formula defining $[G]$), which is undecidable [10]. Luckily, by exploiting the density of the temporal domain, it can be rewritten in the PSPACE-complete fragment $\mathsf{B\overline{B}D}$ (see Section 4), in a very compact way, as follows:

$$\bigwedge_i [G]\big(trip(i) \to (\neg\langle B\rangle trip(i) \wedge \langle\overline{B}\rangle[D][\overline{B}]\neg trip(i))\big) \tag{5}$$

The second requirement states that, for any given bike, a possible new trip must start from the station where its previous trip ended. Since the number of recorded trips of any given bike is obviously finite, this requirement can be easily enforced by the following formula:

$$\bigwedge_{i,j,k} [G]\big(trip(i, j, k) \to \bigwedge_{l \neq k}[A](\langle A\rangle trip(i, l) \to \langle D\rangle trip(i, k))\big) \tag{6}$$

Once again, formula (6) belongs to the undecidable HS fragment $\mathsf{AB\overline{B}D}$ [10]; however, the decidable fragment $\mathsf{B\overline{B}D}$ turns out to be expressive enough to encode such a requirement over dense temporal structures (notice that, over

dense structures, $\langle L \rangle$ is definable using $\langle \overline{B} \rangle$ and $\langle D \rangle$, as shown in [12]):

$$\bigwedge_{i,j,k \neq l,m} [G]\big(trip(i,j,k) \wedge \langle L \rangle trip(i,l,m) \rightarrow$$
$$\langle \overline{B} \rangle(\langle D \rangle \langle \overline{B} \rangle trip(i,k) \wedge [D][\overline{B}](trip(i) \rightarrow trip(i,k)))\big) \quad (7)$$

The basic requirement on table *Maintenance* is the same as for table *Trip*, that is, (a) every interval satisfies at most one repair for each bike, (b) two repairs for the same bike must not overlap, and (c) for any given bike, there exists a non-infinitesimal delay between a repair and the next one:

$$\bigwedge_{i,j,k} [G]\big(repair(i,j,k) \rightarrow \neg \bigvee_{(j,k) \neq (l,m)} repair(i,l,m)\big) \quad (8)$$

$$\bigwedge_{i} disj(repair(i)) \wedge \bigwedge_{i} del(repair(i)) \quad (9)$$

Formula (9) can be rewritten in B$\overline{\text{B}}$D exactly as formula (4).

The next requirement forces a repair to start from a station different from 1 and to end in station 1:

$$\bigwedge_{i,j} [G]\big(\neg repair(i,1,j) \wedge \bigwedge_{k \neq 1} \neg repair(i,j,k)\big) \quad (10)$$

The following requirement constrains maintenance to be performed on a regular basis (infinitely often):

$$\bigwedge_{i} [G]\langle L \rangle repair(i) \quad (11)$$

Requirements on the relationships between trips and repairs state that (i) during maintenance a bike cannot make a trip, and vice versa, and (ii) there is a non-infinitesimal delay between repairs and trips, and vice versa. To express them, we introduce the following two parametric formulae, where φ and ψ are arbitrary propositional formulae:

$$disj(\varphi, \psi) \;=\; [G]\big(\varphi \rightarrow (\neg\psi \wedge \neg\langle B \rangle\psi \wedge$$
$$\neg\langle D \rangle\psi \wedge \neg\langle E \rangle\psi \wedge \neg\langle O \rangle\psi \wedge \neg\langle A \rangle\psi)\big) \quad (12)$$
$$del(\varphi, \psi) \;=\; [G]\big(\varphi \rightarrow \langle A \rangle(\neg\langle D \rangle\psi \wedge \neg\langle E \rangle\psi \wedge \neg\langle O \rangle\psi)\big) \quad (13)$$

Notice that $disj(\varphi, \psi)$ and $del(\varphi, \psi)$ can be viewed as two-variable variants of, respectively, $disj(\varphi)$ and $del(\varphi)$ (we deliberately overloaded the symbols

13

disj and *del* to make such a connection evident), and that $disj(\varphi, \psi)$ differs from $disj(\varphi)$ also for an additional conjunct ($\neg\psi$) in the consequent of $\rightarrow$. Thus, requirements (i) and (ii) above are expressed by the following formula, that can be easily rewritten in $B\overline{B}D$:

$$\bigwedge_i disj(repair(i), trip(i)) \wedge \bigwedge_i disj(trip(i), repair(i)) \wedge$$
$$\bigwedge_i del(repair(i), trip(i)) \wedge \bigwedge_i del(trip(i), repair(i)) \tag{14}$$

where the first (resp, second) line captures requirement (i) (resp., (ii)).

The next requirement states that a repair must start from the end station of the previous trip:

$$\bigwedge_{i,j,k} [G]\Big(trip(i,j,k) \rightarrow \bigwedge_{l \neq k} [A]\big(\langle A\rangle repair(i,l) \rightarrow \bigvee_m \langle D\rangle trip(i,m,l)\big)\Big) \tag{15}$$

Finally, the following requirement forces the first trip after a repair to start from the station where the repair ended:

$$\bigwedge_{i,j,k} [G]\Big(repair(i,j,k) \rightarrow \bigwedge_{l \neq k} [A]\big(\langle A\rangle trip(i,l) \rightarrow \langle D\rangle trip(i,k)\big)\Big) \tag{16}$$

Once more, observe that formulae (15) and (16) belong to an undecidable fragment of HS, but they can be easily rewritten in $B\overline{B}D$.

The above requirements can be viewed as a set of basic rules governing the behavior of the system for the management of bike sharing. We now show that $B\overline{B}D$ can also be exploited to force specific policies involving more advanced temporal constraints like, for instance, those about event duration.

Consider the requirement "two trips of the same bike must be at least t time units apart". Requirements of this kind can be added to an extended scenario where the system also keeps track of users renting bikes and special fares apply for short trips to prevent users from starting a new rental right immediately after the previous one has ended. It can be formalized by making use of a pre-interpreted proposition letter *t-long* which holds true exactly over intervals $[x, y]$ whose length is at least t, that is, $y - x \geq t$.[4]

---

[4]The addition of such a proposition letter does not cause an increase in the complexity

First, we constrain the proposition letter $long\text{-}break_i$ to hold over intervals whose length is at least t (first conjunct of formula (17)) and that are disjoint from trips involving bike i (second and third conjunct of formula (17)) :

$$\bigwedge_i [G](long\text{-}break_i \to t\text{-}long) \land \bigwedge_i disj(long\text{-}break_i, trip(i)) \land \\ \bigwedge_i disj(trip(i), long\text{-}break_i) \quad (17)$$

Formula (17) can be re-written in $B\overline{B}D$ in the usual way.

Then, we force the delay between any pair of trips of the same bike to be equal to at least t time units by the following formula:

$$\bigwedge_i [G]\big(trip(i) \land \langle L \rangle trip(i) \to \langle \overline{B} \rangle (\langle D \rangle long\text{-}break_i \land \neg \langle D \rangle trip(i))\big) \quad (18)$$

Consistency checking. Let us consider the above set of formulae. At a first look, they seem to provide a correct and complete encoding of the intended requirements on the bike sharing system. However, it turns out that they are actually inconsistent. This is because, as it is often the case, an important detail has been neglected: we implemented the specification "a bike trip ends at the station from which the next trip (of the same bike) starts" while the requirements allow for an exception, as in "a bike trip ends at the station from which the next trip (of the same bike) starts, unless the bike goes under maintenance between the two trips". Formally, formula (6) forces every trip to end at the station from which the next trip (of the same bike) starts. Consider the case of two consecutive trips of the same bike i with a repair (of i) in between. By formulae (15) and (16), the first trip ends at some station j, from which the repair starts, and the second trip starts from the station k, where the repair ends. By formula (10), it holds that $j \neq 1$ and $k = 1$, and thus the two consecutive trips of i do not satisfy formula (6).

The decision procedures developed for the HS fragments can be employed to spot inconsistencies like the above one and then to fix the set of requirements. For instance, in the considered case, the inconsistency can be removed

---

of the analysis, as it can be shown that the complexity of the satisfiability problem for resulting extension of $B\overline{B}D$ (over dense linear orders) is still in PSPACE by reducing it to the satisfiability problem for an extension of $B\overline{B}D$ (over dense linear orders), called *cone logic*, which is known to be PSPACE-complete [20].

by rewriting formula (6) in such a way that it takes repairs into account:

$$\bigwedge_{i,j,k}[G]\Big(trip(i,j,k) \to \bigwedge_{l \neq k}[A]\big(\langle A\rangle trip(i,l) \to (\langle D\rangle trip(i,k)$$
$$\vee \langle D\rangle repair(i,k)))\Big) \quad (19)$$

Since all requirements, including the last one, can be expressed in B$\overline{\text{B}}$D, the analysis can be carried out with PSPACE complexity.

4. Complexity of HS fragments: a general overview

HS fragments can be compared to each other in various respects. Here, we are interested in establishing the exact complexity of the satisfiability problem for each (decidable) fragment of HS when interpreted over $\mathbb{Q}$. To this end, we make use of the classification of HS fragments with respect to their relative expressive power [12]. We say that HS modality $\langle X\rangle$ is *definable* in a fragment $\mathcal{F}$ (denoted $\langle X\rangle \lhd \mathcal{F}$) if, for any proposition letter $p \in \mathcal{AP}$, the formula

$$\langle X\rangle p \leftrightarrow \psi$$

is valid in the linear order of $\mathbb{Q}$ (we denote it by $\langle X\rangle p \equiv \psi$), for some $\mathcal{F}$-formula ψ over p. The equivalence $\langle X\rangle p \equiv \psi$ is called a *definability equation* (or simply *definability*) for $\langle X\rangle$ in $\mathcal{F}$. By identifying the complete set of definabilities that hold in the linear order of $\mathbb{Q}$, we are able to characterize the subset of *expressively different* fragments of HS (for this particular class of models). These are the fragments that we classify in terms of the complexity of their satisfiability problem.

In [12], Aceto et al. proved that there are precisely 18 essential definabilities, that is, such that no one can be derived from the others, that hold among HS modalities over dense linear orders (equivalently $\mathbb{Q}$):

1. $\langle L\rangle p \equiv \langle A\rangle\langle A\rangle p$ and its mirror image;
2. $\langle L\rangle p \equiv \langle \overline{B}\rangle[E]\langle \overline{B}\rangle\langle E\rangle p$ and its mirror image;
3. $\langle L\rangle p \equiv \langle O\rangle(\langle O\rangle\top \wedge [O]\langle D\rangle\langle O\rangle p)$ and its mirror image;
4. $\langle L\rangle p \equiv \langle \overline{B}\rangle[D]\langle \overline{B}\rangle\langle D\rangle\langle \overline{B}\rangle p$ and its mirror image;
5. $\langle L\rangle p \equiv \langle O\rangle[E]\langle O\rangle\langle O\rangle p$ and its mirror image;
6. $\langle L\rangle p \equiv \langle O\rangle(\langle O\rangle\top \wedge [O]\langle B\rangle\langle O\rangle\langle O\rangle p)$ and its mirror image;
7. $\langle L\rangle p \equiv \langle O\rangle(\langle O\rangle\top \wedge [O][\overline{L}]\langle O\rangle\langle O\rangle p)$ and its mirror image;
8. $\langle O\rangle p \equiv \langle E\rangle\langle \overline{B}\rangle p$ and its mirror image;
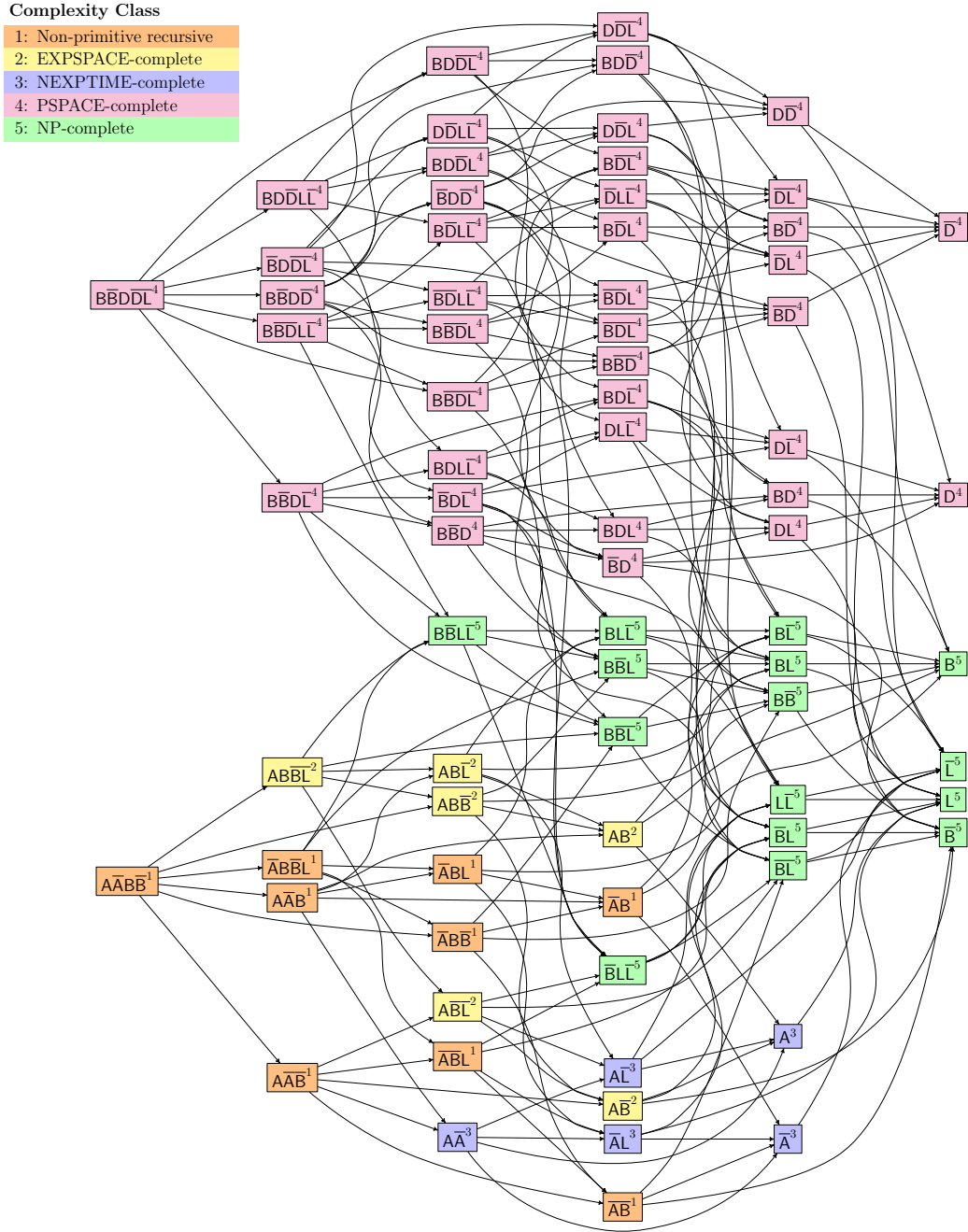
16

Figure 2: Decidable HS fragments over the linear order of the rational numbers: their relative expressive power and their computational complexity.

9. $\langle D \rangle p \equiv \langle E \rangle \langle B \rangle p$;
10. $\langle \overline{D} \rangle p \equiv \langle \overline{E} \rangle \langle \overline{B} \rangle p$.

It is worth pointing out that while proving that a given definability is correct is relatively easy, showing that a given set of definabilities is complete is quite hard, as it requires a complex analysis based on bisimulations or bisimulation-based games [12]. The set of all the essential definabilities allows us to draw a graph of expressively different HS fragments, which, in the case of dense linear orders, contains 966 of the 4095 different non-empty subsets of 12 modalities (we do not consider the empty set of modalities, corresponding to propositional logic). Of these, precisely 130 are decidable [15, 16, 17, 18, 19, 20].[5] They are shown in Fig. 2. Notice that we do not depict decidable fragments containing modalities $\langle E \rangle$ or $\langle \overline{E} \rangle$, as their mirror images, containing, respectively, $\langle B \rangle$ or $\langle \overline{B} \rangle$ instead, are depicted.

Undecidable fragments. As shown in [10], a minimal condition for an HS fragment to be undecidable over $\mathbb{Q}$ is to contain, as definable, the fragments O, AD, or A$\overline{\text{D}}$ (or one of the mirror fragments). Each of these fragments is indeed expressive enough to make it possible to reduce the *finite tiling problem* [26] to it. In addition, it is well known from [14] that fragments containing, as definable, modality $\langle B \rangle$ (*begins*) or $\langle \overline{B} \rangle$ (*begun by*) together with modality $\langle E \rangle$ (*ends*) or $\langle \overline{E} \rangle$ (*ended by*) are undecidable.

Non-primitive recursive fragments. The decidability of A$\overline{\text{A}}$B$\overline{\text{B}}$ over $\mathbb{Q}$ has been proved in [15]. The proof is based on an algorithm of non-primitive recursive complexity. This leaves open the problem of establishing the exact complexity of (most of the fragments of) A$\overline{\text{A}}$B$\overline{\text{B}}$. In Section 7, we solve it by proving the hardness of the fragments $\overline{\text{A}}$B and $\overline{\text{A}}\overline{\text{B}}$ (and thus of each fragment containing one of them, as definable) for the non-primitive recursive class.

ExpSpace-complete fragments. As a consequence of the results given in [16], we know that AB$\overline{\text{B}}\overline{\text{L}}$ is in ExpSpace. In Section 6, we show that each fragment containing AB or A$\overline{\text{B}}$ is ExpSpace-hard over $\mathbb{Q}$.

It is worth pointing out the different positioning of the HS fragments $\overline{\text{A}}$B and $\overline{\text{A}}\overline{\text{B}}$ with respect to the fragments AB and A$\overline{\text{B}}$. One might indeed expect the former to behave like the latter, while this is definitely not the case, as they ended up in quite different complexity classes (non-primitive recursive

_____

[5]We would like to notice that in [1] we wrote 146 instead of 130 as we mistakenly considered as semantically different some HS fragments which, in view of the essential definabilities above, are in fact equivalent.

vs. ExpSpace-complete). In fact, this does not come as a surprise, since the modalities of the first pair of HS fragments support full access to intervals in the past and, yet, make it possible to predicate about intervals ending in the future, while those of the second pair cannot access intervals starting or ending strictly before the current one.

NExpTime-complete fragments. The collection of NExpTime-complete fragments of HS includes those fragments that have been the first to be proved decidable. These results can be found in [17].

PSpace-complete fragments. PSpace-complete fragments have been identified in [18, 19, 20], where it has been proved that $B\overline{B}D\overline{D}L\overline{L}$ is PSpace-complete, and that a sufficient condition for any fragment of HS to be PSpace-hard is to contain, as definable, the fragments D or $\overline{D}$.

NP-complete fragments. The results for the PSpace-complete fragments leave open the problem of establishing the complexity of those fragments of $B\overline{B}D\overline{D}L\overline{L}$ containing neither D nor $\overline{D}$. Among them, the most expressive one is $B\overline{B}L\overline{L}$. In Section 5, we prove that $B\overline{B}L\overline{L}$ and all of its fragments are decidable in NP (NP-hardness trivially holds).

5. NP-complete fragments

In this section, we show that $B\overline{B}L\overline{L}$ is in NP (NP-completeness immediately follows, as propositional logic is a syntactic fragment of $B\overline{B}L\overline{L}$).

The proof of NP-membership consists of two main steps. The first step reduces the satisfiability of a formula φ to the existence of a suitable pseudo-model, that is, it shows that φ is satisfiable if and only if there exists a corresponding, finitely-representable structure that satisfies a given set of constraints induced by φ. The second step proves that (i) each satisfiable formula admits a pseudo-model of size at most $P(|\varphi|)$, where P is a polynomial, and (ii) constraint satisfaction (by the pseudo-model) can be checked in nondeterministic polynomial time.

The proof of item (i) makes use of two fundamental ingredients. The first one is a fairly natural spatial representation of intervals, originally exploited by Venema in [27], which maps intervals into points and, for any such point and any Allen relation, provides a spatial characterization of all the points which are in such a relation with it, that is, it identifies the region of the plane all these points belong to. The second one is the monotonicity of modalities $\langle L \rangle / \langle \overline{L} \rangle$ and $\langle B \rangle / \langle \overline{B} \rangle$. Points (equivalently, the intervals they represent)

are first partitioned on the basis of their first coordinate and then clustered according to the $\langle L \rangle / \langle \overline{L} \rangle$ requests of their second coordinate. Thanks to the monotonicity of modalities $\langle L \rangle / \langle \overline{L} \rangle$, these clusters can be represented using polynomial space (in the size of the input formula $|\varphi|$). A point belonging to any such cluster either is a unique witness in the model (a singleton) or can be merged with its neighbours into a maximal open segment. Next, a new set of clusters is associated with each point of the rational line, being it a singleton or an element of an open segment, on the basis of its $\langle B \rangle / \langle \overline{B} \rangle$ requests. Thanks to the maximality of open segments and the monotonicity of modalities $\langle B \rangle / \langle \overline{B} \rangle$, the set of sets of clusters associated with the elements of an open segment can be represented in polynomial space. Finally, despite the fact that the number of possible sets of clusters that may label an open segment is, in principle, exponential in $|\varphi|$, we show that if φ has a model in $\mathbb{Q}$, then there exists a pseudo-model for φ that features only a polynomial number of them.

To make the proof easier to follow, we shall first show that the sub-fragment $\mathsf{L}\overline{\mathsf{L}}$ belongs to NP, and then we shall generalize the proof to $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$.

5.1. Preliminaries

In the following, we recall terminology and notation which are common in the (interval) temporal logic setting. We define the *closure* of a formula φ, denoted by $\mathcal{C}l(\varphi)$, as the set of all sub-formulae of φ and all their negations (we identify $\neg\neg\psi$ with ψ, $\neg\langle B \rangle\psi$ with $[B]\neg\psi$, $\neg[B]\psi$ with $\langle B \rangle\neg\psi$, and so on). For technical reasons that will become clear soon, we also introduce the *extended closure* of φ, denoted by $\mathcal{E}\mathcal{C}l(\varphi)$, which, in addition, includes all formulae of the form $\langle X \rangle\psi$ and $[X]\psi$, for $X \in \{B, \overline{B}, L, \overline{L}\}$ and $\psi \in \mathcal{C}l(\varphi)$. We define a φ-*atom* F (atom F for short) as a maximal, locally consistent subset of $\mathcal{E}\mathcal{C}l(\varphi)$. Formally, F is a subset of $\mathcal{E}\mathcal{C}l(\varphi)$ such that (i) for all $\psi \in \mathcal{E}\mathcal{C}l(\varphi)$, $\psi \in F$ if and only if $\neg\psi \notin F$, and (ii) for all $\psi = \psi_1 \vee \psi_2 \in \mathcal{E}\mathcal{C}l(\varphi)$, $\psi \in F$ if and only if $\psi_1 \in F$ or $\psi_2 \in F$. We denote the set of all φ-atoms by *atoms*(φ). We say that an atom F is B-reflexive if and only if (i) for each $[B]\psi \in F$, $\psi \in F$, and (ii) for each $[\overline{B}]\psi \in F$, $\psi \in F$.

Given an atom F and $X \in \{B, \overline{B}, L, \overline{L}\}$, let $\mathcal{R}eq_X(F)$ be the set of *requests* of F for the relation R_X, namely, the set of formulae $\psi \in \mathcal{C}l(\varphi)$ such that $\langle X \rangle\psi \in F$. Similarly, let $\mathcal{O}bs(F) = F \cap \mathcal{C}l(\varphi)$ be the set of *observables* of F—intuitively, the observables of F are those formulae $\psi \in F$, belonging to $\mathcal{C}l(\varphi)$, that fulfil requests of the form $\langle X \rangle\psi$ associated with other atoms. Given a $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$ formula φ, we define a φ-*labeled interval structure* for

φ (*LIS* for short) as a pair $\mathbf{L} = (\mathbb{Q}, \mathcal{L})$, where $\mathcal{L} : \mathbb{I}(\mathbb{Q}) \to atoms(\varphi)$ is a function that maps each interval on $\mathbb{Q}$ in an atom such that, for every pair of intervals $[x, y], [x', y'] \in \mathbb{I}(\mathbb{Q})$ and each $X \in \{B, \overline{B}, L, \overline{L}\}$, if $[x, y]R_X[x', y']$, then $\mathcal{O}bs(\mathcal{L}([x', y'])) \subseteq \mathcal{R}eq_X(\mathcal{L}([x, y]))$.

It can be easily checked that *LIS*s satisfy universal (temporal) conditions by construction. First, we observe that if $[X]\psi$ is a subformula of φ, then $\langle X \rangle \psi$ belongs to $\mathcal{ECl}(\varphi)$. Now, let $[x, y] \in \mathbb{I}(\mathbb{Q})$ and $\mathcal{L}([x, y]) = F$, for some atom F. If $[X]\psi \in F$, then, by definition of atom, $\langle X \rangle \neg \psi \notin F$, and thus $\neg \psi$ does not belong to $\mathcal{R}eq_X(F)$. By definition of *LIS*, it immediately follows that for all intervals $[x', y'] \in \mathbb{I}(\mathbb{Q})$ such that $[x, y]R_X[x', y']$, $\neg \psi \notin \mathcal{O}bs(\mathcal{L}([x', y']))$, and thus, again by definition of atom, $\psi \in \mathcal{O}bs(\mathcal{L}([x', y']))$.

We say that $\mathbf{L}$ is a *fulfilling LIS* for φ if it satisfies the following properties: *(i)* for every interval $[x, y] \in \mathbb{I}(\mathbb{Q})$, each $X \in \{B, \overline{B}, L, \overline{L}\}$, and each $\psi \in \mathcal{R}eq_X(\mathcal{L}([x, y]))$, there exists $[x', y'] \in \mathbb{I}(\mathbb{Q})$ such that $[x, y]R_X[x', y']$ and $\psi \in \mathcal{O}bs(\mathcal{L}([x', y']))$, and *(ii)* there exists an interval $[x, y] \in \mathbb{I}(\mathbb{Q})$ such that $\varphi \in \mathcal{L}([x, y])$. The next theorem, whose proof is straightforward (see, e.g., [17]), states that fulfilling *LIS*s suffice to witness satisfiability.

Theorem 1. *Let φ be a $\mathsf{B\overline{B}L\overline{L}}$ formula. It is satisfiable over $\mathbb{Q}$ if and only if there exists a fulfilling LIS $\mathbf{L} = (\mathbb{Q}, \mathcal{L})$ for it.*

5.2. The sub-fragment $\mathsf{L\overline{L}}$

To start with, we observe that, by the semantics of modality $\langle L \rangle$, intervals with the same right endpoint agree on the truth of formulae of the form $\langle L \rangle \varphi$ ($\langle L \rangle$-formulae for short). Symmetrically, intervals with the same left endpoint agree on $\langle \overline{L} \rangle$-formulae. Given an interval model $M = \langle \mathbb{I}(\mathbb{Q}), V \rangle$ for a formula φ and a rational number x, we define the $\mathsf{L\overline{L}}$-*requests* of x as the pair of sets $(L_x, \overline{L}_x)$, where L_x contains all formulae $\psi \in \mathcal{Cl}(\varphi)$ such that $\langle L \rangle \psi$ is true over all intervals $[y, x]$, with $y < x$ (L-*requests*), and $\overline{L}_x$ contains all formulae $\psi \in \mathcal{Cl}(\varphi)$ such that $\langle \overline{L} \rangle \psi$ is true over all intervals $[x, y]$, with $x < y$ ($\overline{\mathsf{L}}$-*requests*). By the transitivity of $\langle L \rangle$ (resp., $\langle \overline{L} \rangle$), we have that the set of L-*requests* (resp., $\overline{\mathsf{L}}$-*requests*) is monotone with respect to $<$, that is, for every pair of rational numbers x, y, with $x < y$, it holds that $L_x \supseteq L_y$ (resp., $\overline{L}_x \subseteq \overline{L}_y$).

Since the closure of a formula is a finite set, we can partition $\mathbb{Q}$ into a finite number of clusters whose elements have the same $\mathsf{L\overline{L}}$-requests ($\mathsf{L\overline{L}}$-*clusters*). It is possible to define the partition in such a way that each cluster is either a singleton (*point-interval*) or a convex, open set (*open interval*).

21

Definition 1. *An* $\mathsf{L}\overline{\mathsf{L}}$*-sequence is a finite sequence of triples of the form:*

$$(L_1, \overline{L}_1, Type_1), (L_2, \overline{L}_2, Type_2), \ldots, (L_n, \overline{L}_n, Type_n),$$

where, for $1 \leq i \leq n$, L_i *and* $\overline{L}_i$ *are sets of* L*-requests and* $\overline{\mathsf{L}}$*-requests, respectively, and* $Type_i$ *is either* point *(for a point-interval) or* interval *(for an open interval), such that:*

(i) $L_i \supseteq L_{i+1}$ *and* $\overline{L}_i \subseteq \overline{L}_{i+1}$, *for* $1 \leq i < n$ *(monotonicity);*
(ii) *for all* i, j, *with* $1 \leq i < j \leq n$, *if* $Type_i = Type_j =$ interval, *then* $(L_i, \overline{L}_i) \neq (L_j, \overline{L}_j)$ *(maximality);*
(iii) $Type_1 = Type_n =$ interval *(the first and last clusters are open sets);*
(iv) *for each* $1 < i < m$, *if* $Type_i =$ point, *then* $Type_{i-1} = Type_{i+1} =$ interval *(no consecutive clusters of type* point*).*

Since both $|L_i|$ and $|\overline{L}_i|$ are less than or equal to $|\mathcal{Cl}(\varphi)| = 2 \cdot |\varphi|$, for each $1 \leq i \leq n$, there are at most $8 \cdot |\varphi| - 1$ distinct triples in the sequence: by (i) and (ii), there are at most $4 \cdot |\varphi|$ triples of type *interval*, and, by (iii) and (iv), there are at most $4 \cdot |\varphi| - 1$ triples of type *point*.

Definition 2. *A pseudo-model for an* $\mathsf{L}\overline{\mathsf{L}}$ *formula* φ *is a labeled* $\mathsf{L}\overline{\mathsf{L}}$*-sequence* $(L_1, \overline{L}_1, Type_1), \ldots, (L_n, \overline{L}_n, Type_n)$, *where the labeling function* $f : \{(i,j) : 1 \leq i \leq j \leq n\} \mapsto 2^{atoms(\varphi)}$ *satisfies the following constraints on the cardinality of the images:*

- $|f(i,i)| = 0$ *if* $Type_i =$ point *(we assumed the strict semantics);*
- $|f(i,j)| = 1$ *if* $i \neq j$ *and* $Type_i = Type_j =$ point*;*
- $|f(i,j)| \geq 1$ *if* $Type_i =$ interval *or* $Type_j =$ interval *(or both),*

as well as the following featuring, consistency, and fulfillingness conditions:

- *there exist a pair* (i,j), *with* $i < j$, *and an atom* $F \in f(i,j)$ *such that* $\varphi \in F$ *(featuring condition);*
- *for each pair* (i,j), *with* $i \leq j$, *and each atom* $F \in f(i,j)$, $\mathcal{Req}_{\overline{L}}(F) = \overline{L}_i$, $\mathcal{Req}_L(F) = L_j$, *if* $Type_j =$ point *(resp., interval), then* $\mathcal{Obs}(F) \subseteq \overline{L}_{j+1}$ *(resp.,* $\mathcal{Obs}(F) \subseteq \overline{L}_j$*), and if* $Type_i =$ point *(resp., interval), then* $\mathcal{Obs}(F) \subseteq L_{i-1}$ *(resp.,* $\mathcal{Obs}(F) \subseteq L_i$*) (consistency);*
- *for each index* i, *with* $1 \leq i \leq n$, *and each formula* $\psi \in L_i$, *if* $Type_i =$ point *(resp., interval), there exist a pair* (j, j'), *with* $i < j \leq j'$ *(resp.,* $i \leq j \leq j'$*), and an atom* $F \in f(j, j')$ *such that* $\psi \in F$ *(*L*-fulfillingness);*
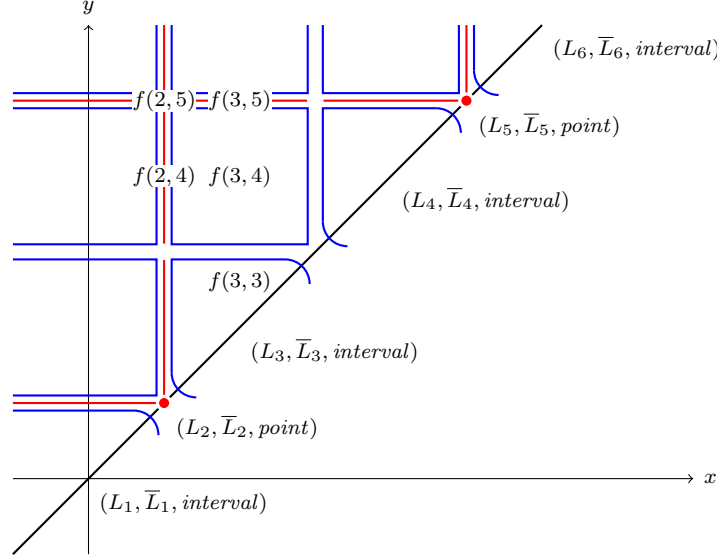- $\overline{\mathsf{L}}$*-fulfillingness is defined analogously.*

Figure 3: A graphical account of an $\mathsf{L\overline{L}}$ pseudo-model.

The consistency condition takes care of $[L]$- and $[\overline{L}]$-formulae, while the fulfillment conditions guarantee that $\langle L\rangle$- and $\langle\overline{L}\rangle$-formulae are satisfied. We call the labeling function f an $\mathsf{L\overline{L}}$-*witness* for the $\mathsf{L\overline{L}}$-sequence $(L_1, \overline{L}_1, Type_1), \ldots, (L_n, \overline{L}_n, Type_n)$. Figure 3 gives a graphical account of an $\mathsf{L\overline{L}}$ pseudo-model, that features 6 clusters, by exploiting the natural correspondence between intervals $[x, y]$ and points (x, y) on the cartesian half-plane $x \leq y$ of $\mathbb{Q} \times \mathbb{Q}$ [27]. Clusters of type *point* (2 and 5, in red) and *interval* (1, 3, 4, and 6, in blue) partition the "half plane" into regions of different shape, that are then labeled by the function f. Such regions can be either "open rectangles", e.g., $f(3, 4)$, "open triangles", e.g., $f(3, 3)$, "open segments", e.g., $f(2, 4)$ and $f(3, 5)$, or "points", e.g., $f(2, 5)$. Cardinality, featuring, consistency, and fulfillingness conditions guarantee that such a partitioning respects the semantics of the formulae and the shape of the different regions.

The next lemma proves that the satisfiability of an $\mathsf{L\overline{L}}$ formula φ can be reduced to the existence of an $\mathsf{L\overline{L}}$ pseudo-model for it.

Lemma 1. *An $\mathsf{L\overline{L}}$ formula φ is satisfiable over $\mathbb{Q}$ if and only if there exists an $\mathsf{L\overline{L}}$ pseudo-model for it.*

Proof. We first prove the left-to-right implication. Let $\mathbf{L} = (\mathbb{Q}, \mathcal{L})$ be a fulfilling *LIS* for φ, whose existence is guaranteed by Theorem 1. We build

an $\mathsf{L\overline{L}}$ pseudo-model for φ as follows. Let $[\![x_1]\!], [\![x_2]\!], \ldots, [\![x_n]\!]$ be a partition of $\mathbb{Q}$ that satisfies the following properties: *(i)* the representatives are totally ordered, that is, $x_1 < x_2 < \ldots < x_n$; *(ii)* each $[\![x_i]\!]$ is either the singleton $\{x_i\}$ or an open interval (x_i^-, x_i^+) of $\mathbb{Q}$ (notice that x_i^-, resp., x_i^+, might be $-\infty$, resp., $+\infty$); *(iii)* for $i = 1, \ldots, n$, all elements in $[\![x_i]\!]$ have the same $\mathsf{L\overline{L}}$-requests $(L_i, \overline{L}_i)$; *(iv)* for each pair of open intervals $[\![x_i]\!], [\![x_j]\!]$, with $i \neq j$, $(L_i, \overline{L}_i) \neq (L_j, \overline{L}_j)$.

Since the set of L-*requests* (resp., $\overline{\mathsf{L}}$-*requests*) is monotone with respect to $<$ and there is a finite number of possible (different) $\mathsf{L\overline{L}}$-*requests*, such a partition can always be defined. Moreover, the first and the last element of the partition $[\![x_1]\!]$ and $[\![x_n]\!]$ are necessarily open intervals and, since $\mathbb{Q}$ is dense, there are not two consecutive singletons. We define the $\mathsf{L\overline{L}}$-sequence of the $\mathsf{L\overline{L}}$ pseudo-model for φ as follows:

$$\sigma_{L\bar{L}} = (L_1, \overline{L}_1, \mathit{Type}_1), (L_2, \overline{L}_2, \mathit{Type}_2), \ldots, (L_n, \overline{L}_n, \mathit{Type}_n),$$

where, for $i = 1, \ldots, n$, $(L_i, \overline{L}_i)$ are the $\mathsf{L\overline{L}}$-requests of $[\![x_i]\!]$ and $\mathit{Type}_i = \mathit{point}$ (resp., $\mathit{Type}_i = \mathit{interval}$) if $[\![x_i]\!]$ is a singleton (resp., an open interval). It can be easily checked that $\sigma_{L\bar{L}}$ is indeed an $\mathsf{L\overline{L}}$-sequence. The labeling function f of the $\mathsf{L\overline{L}}$ pseudo-model can be defined as follows:

$$f(i, j) = \{\mathcal{L}([x, y]) \mid x \in [\![x_i]\!] \text{ and } y \in [\![x_j]\!]\}.$$

In the following, we show that f satisfies cardinality, featuring, consistency, and fulfillingness constraints.

The satisfaction of cardinality constraints immediately follows from the definition of f and the properties of the partition: *(i)* if $[\![x_i]\!]$ is a singleton, then the interval $[x_i, x_i]$ does not belong to $\mathbb{I}(\mathbb{Q})$ (strict semantics), and thus $f(i, i)$ is the empty set; *(ii)* if both $[\![x_i]\!]$ and $[\![x_j]\!]$ $(i < j)$ are singletons, then $f(i, j)$ contains the labeling of the interval $[x_i, x_j]$ only; *(iii)* in all the other cases $f(i, j)$ contains at least one atom.

As for the featuring condition, $\mathbf{L} = (\mathbb{Q}, \mathcal{L})$ is a fulfilling *LIS* for φ and thus there exists an interval $[x, y]$ such that $\varphi \in \mathcal{L}([x, y])$. Let $[\![x_i]\!]$ and $[\![x_j]\!]$ be the two elements in the partition such that $x \in [\![x_i]\!]$ and $y \in [\![x_j]\!]$. By the definition of f, $\mathcal{L}([x, y]) \in f(i, j)$ and thus the featuring condition is fulfilled.

We now check that the consistency conditions are satisfied as well. Let $F \in f(i, j)$. Since all elements in $[\![x_j]\!]$ (resp., $[\![x_i]\!]$) have the same set of L-requests (resp., $\overline{\mathsf{L}}$-requests), it immediately follows that $\mathcal{R}eq_L(F) = L_j$ (resp., $\mathcal{R}eq_{\overline{L}}(F) = \overline{L}_i$). Suppose now that $[\![x_j]\!]$ is the open interval (x_j^-, x_j^+). By the

definition of f, we can find an interval $[x, y]$ such that $x \in [\![x_i]\!]$, $y \in [\![x_j]\!]$, and $F = \mathcal{L}([x, y])$. Since $\mathbb{Q}$ is dense, we can find an interval $[x', y']$ with $y < x' < x_j^+$. By the definition of LIS, it holds that $\mathcal{O}bs(\mathcal{L}[x, y]) \subseteq \mathcal{R}eq_{\overline{L}}(\mathcal{L}([x', y']))$, and thus $\mathcal{O}bs(F) \subseteq \overline{L}_j$. The other cases, as well as the last consistency requirement, can be dealt with in the very same way.

To prove that the condition of L-fulfillingness is satisfied, consider a formula $\psi \in L_i$ for some $1 \leq i \leq n$. Since $\mathbf{L} = (\mathbb{Q}, \mathcal{L})$ is a fulfilling LIS, there is an interval $[x, y]$, with $x_i < x$, such that $\psi \in \mathcal{L}([x, y])$. Now, let $x \in [\![x_j]\!]$ and $y \in [\![x_{j'}]\!]$. By the definition of f, $\mathcal{L}([x, y]) \in f(j, j')$, and, by the definition of the partition, $j \leq j'$ and $i < j$, if $[\![x_i]\!]$ is a singleton, and $i \leq j$, if $[\![x_i]\!]$ is an open interval. $\overline{\mathsf{L}}$-fulfillingness can be proved in a similar way.

To prove the right-to-left implication, let us consider an $\mathsf{L}\overline{\mathsf{L}}$ pseudo-model consisting of the $\mathsf{L}\overline{\mathsf{L}}$-sequence $(L_1, \overline{L}_1, Type_1), \ldots, (L_n, \overline{L}_n, Type_n)$ for φ and the associated labeling function f, that satisfies the cardinality, featuring, consistency, and fulfillingness conditions. The fulfilling LIS $\mathbf{L} = (\mathbb{Q}, \mathcal{L})$ satisfying φ is obtained as the limit of an infinite sequence of finite LISs $\mathbf{L}_0 = (\mathbb{D}_0, \mathcal{L}_0), \mathbf{L}_1 = (\mathbb{D}_1, \mathcal{L}_1), \mathbf{L}_2 = (\mathbb{D}_2, \mathcal{L}_2), \ldots$. Each finite LIS $\mathbf{L}_i = (\mathbb{D}_i, \mathcal{L}_i)$, with $\mathbb{D}_i = \langle D_i, < \rangle$, is paired with a clustering function $Clst_i : D_i \mapsto \{1, \ldots, n\}$ such that, for each interval $[x, y] \in \mathbb{I}(\mathbb{D}_i)$, $\mathcal{L}_i([x, y]) \in f(Clst_i(x), Clst_i(y))$. By the featuring condition, there exist a pair of indexes j, k, with $j \leq k$, and an atom $F_\varphi \in f(j, k)$ such that $\varphi \in F_\varphi$.

The initial LIS $\mathbf{L}_0 = (\mathbb{D}_0, \mathcal{L}_0)$, with $\mathbb{D}_0 = \langle D_0, < \rangle$, is defined as follows: we let $D_0 = \{1, 2, \ldots, n\}$, $\mathcal{L}_0([j, k]) = F_\varphi$, $\mathcal{L}([l, m])$ equal to an arbitrary atom in $f(l, m)$, for all $(l, m) \neq (j, k)$, and $Clst_0(h) = h$, for all $1 \leq h \leq n$. Then, the LIS $\mathbf{L}_{i+1}$ is obtained from the LIS $\mathbf{L}_i$ by selecting one of its "defect" and fixing it (as it will soon be clear, for all i, we can find a defect in $\mathbf{L}_i$).

Five distinct types of defect must be taken into consideration (in the following, we enumerate the elements of D_i as $x_1 < x_2 < \ldots < x_{|D_i|}$).

The first two types of defect are L- and $\overline{\mathsf{L}}$-*requests* which are not fulfilled.

L-defects: an L-defect occurs if there exist an interval $[x_a, x_b] \in \mathbb{I}(\mathbb{D}_i)$ and a formula $\langle L \rangle \psi \in \mathcal{L}_i([x_a, x_b])$ such that for all $[x_c, x_d] \in \mathbb{I}(\mathbb{D}_i)$, if $x_c > x_b$, then $\psi \notin \mathcal{L}_i([x_c, x_d])$.

To fix such a defect, we must insert in $\mathbf{L}_{i+1}$ a new interval that satisfies ψ. To find its correct placement, we proceed as follows. Let $j = Clst_i(x_b)$. By the definition of $Clst_i$ and the consistency of f, $\psi \in L_j = \mathcal{R}eq_L(\mathcal{L}_i([x_a, x_b]))$ (ψ belongs to the set of L-requests of the j-th cluster). Hence, by the L-fulfillingness condition, there exist two indexes k, l, with $j \leq k \leq l \leq n$, and

an atom F_ψ such that $\psi \in F_\psi$ and $F_\psi \in f(k, l)$. Let x_c and x_d be the greatest elements in $\mathbb{D}_i$ such that $Clst_i(x_c) = k$ and $Clst_i(x_d) = l$, and let x'_c and x'_d be the immediate successors of x_c and x_d in $\mathbb{D}_i$, respectively. The endpoints of the new interval $[x', y']$ are defined as follows:

$$
x' = \begin{cases}
x_c & \text{if } Type_k = \text{point}; \\
x_c + 1 & \text{if } Type_k = \text{interval and } x_c = \max(\mathbb{D}_i); \\
\frac{x_c + x'_c}{2} & \text{otherwise},
\end{cases}
$$

$$
y' = \begin{cases}
x_d & \text{if } Type_l = \text{point}; \\
x' + 1 & \text{if } x_c = x_d \text{ and } x_d = \max(\mathbb{D}_i); \\
\frac{x' + x'_d}{2} & \text{if } x_c = x_d \text{ and } x_d \neq \max(\mathbb{D}_i); \\
x_d + 1 & \text{if } x_c \neq x_d \text{ and } x_d = \max(\mathbb{D}_i); \\
\frac{x_d + x'_d}{2} & \text{otherwise}.
\end{cases}
$$

Hence, we let $D_{i+1} = D_i \cup \{x', y'\}$ and we define the labeling $\mathcal{L}_{i+1}$ as:

$$
\mathcal{L}_{i+1}([x, y]) = \begin{cases}
F_\psi & \text{if } x = x' \text{ and } y = y'; \\
F \in f(k, Clst_i(y)) & \text{if } x = x', y \neq y', \text{ and } x' \notin D_i; \\
F \in f(Clst_i(x), l) & \text{if } x \neq x', y = y', \text{ and } y' \notin D_i; \\
F \in f(Clst_i(x), k) & \text{if } y = x' \text{ and } x' \notin D_i; \\
F \in f(l, Clst_i(y)) & \text{if } x = y' \text{ and } y' \notin D_i; \\
\mathcal{L}_i([x, y]) & \text{otherwise}.
\end{cases}
$$

As for $Clst_{i+1}$, we let $Clst_{i+1}(x') = k$, $Clst_{i+1}(y') = l$, and $Clst_{i+1}(x) = Clst_i(x)$ for all $x \in D_i$.

$\overline{\mathsf{L}}$**-defects:** an $\overline{\mathsf{L}}$-defect occurs if there exist an interval $[x_a, x_b] \in \mathbb{I}(\mathbb{D}_i)$ and a formula $\langle \overline{L} \rangle \psi \in \mathcal{L}([x_a, x_b])$ such that for all $[x_c, x_d] \in \mathbb{I}(\mathbb{D}_i)$, if $x_d < x_a$, then $\psi \notin \mathcal{L}([x_c, x_d])$. To fix such a defect, we must insert in $\mathbf{L}_{i+1}$ a new interval that satisfies ψ. We can proceed exactly as in the previous case.

The remaining three types of defects are: $\mathbb{D}_i$ is not dense, $\mathbb{D}_i$ has a minimum, and $\mathbb{D}_i$ has a maximum. They occur in every $\mathbf{L}_i$ and can be fixed as follows.

Disc-defects: we pick two consecutive elements x_a and x_{a+1} in D_i and we add either one or two elements to D_{i+1}, depending on the $Type$ of the j-th and the k-th triple, where $j = Clst_i(x_a)$ and $k = Clst_i(x_{a+1})$:

- if $Type_j = point$, then we add the point $x' = \frac{x_a + x_{a+1}}{2}$ to D_{i+1} and we put $Clst_{i+1}(x') = k$;
- if $Type_k = point$, then we add the point $x' = \frac{x_a + x_{a+1}}{2}$ to D_{i+1} and we put $Clst_{i+1}(x') = j$;
- if $Type_j = Type_k = interval$, then we add the points $x' = \frac{2x_a + x_{a+1}}{3}$ and $y' = \frac{x_a + 2x_{a+1}}{3}$ to D_{i+1}, putting $Clst_{i+1}(x') = j$ and $Clst_{i+1}(y') = k$, respectively.

It is worth pointing out that it cannot be the case that $Type_j = Type_k = point$: the initial LIS has a witness for each triple and thus either $k = j$ or $k = j + 1$ and there cannot be two consecutive triples of type $point$. The labeling function $\mathcal{L}_{i+1}$ extends $\mathcal{L}_i$ by putting $\mathcal{L}_{i+1}([x,y]) = F$, for some $F \in f(Clst_{i+1}(x), Clst_{i+1}(y))$, for all intervals $[x,y] \notin \mathbb{I}(\mathbb{D}_i)$.

Min-defects: given $x_1 = \min(\mathbb{D}_i)$, we add the element $x' = x_1 - 1$ to D_{i+1}, and we let $Clst_{i+1}(x') = 1$. For each element $y > x'$, we choose an atom $F \in f(Clst_{i+1}(x'), Clst_i(y))$ and we put $\mathcal{L}_{i+1}([x',y]) = F$.

Max-defects: given $x_{|D_i|} = \max(\mathbb{D}_i)$, we add the element $x' = x_{|D_i|} + 1$ to D_{i+1}, and we let $Clst_{i+1}(x') = n$. For each element $x < x'$, we choose an atom $F \in f(Clst_i(x), Clst_{i+1}(x'))$ and we put $\mathcal{L}_{i+1}([x,x']) = F$.

By fixing all the defects according to a fair strategy which guarantees that no $\langle L \rangle$-/$\langle \overline{L} \rangle$-request is indefinitely postponed and the final domain is dense, we can force the limit structure $\mathbf{L}_\omega = \bigcup_{i \in \mathbb{N}} \mathbf{L}_i$ to be a fulfilling LIS for φ. $\square$

We complete the proof by showing that the existence of a pseudo-model can be checked in polynomial time. We already observed that the number of clusters, and, thus, the length of an $\mathsf{L}\overline{\mathsf{L}}$-sequence, is bounded by $8 \cdot |\varphi| - 1$. Then, the size of the domain of the labeling function f is at most $\frac{(8 \cdot |\varphi| - 1)(8 \cdot |\varphi|)}{2}$. However, each element in the image of f is a subset of $atoms(\varphi)$ whose size might be exponential. The next lemma states the existence of polynomially bounded $\mathsf{L}\overline{\mathsf{L}}$-witness for $\mathsf{L}\overline{\mathsf{L}}$ formulae which are satisfiable over $\mathbb{Q}$.

Lemma 2. *Let φ be an $\mathsf{L}\overline{\mathsf{L}}$ formula which is satisfiable over $\mathbb{Q}$. For every pseudo-model $(L_1, \overline{L}_1, Type_1), \ldots, (L_n, \overline{L}_n, Type_n)$ for φ, there is an $\mathsf{L}\overline{\mathsf{L}}$-witness f such that for all i, j, with $i \leq j$, $|f(i,j)| \leq 2 \cdot |\varphi|$.*

Proof. Let $(L_1, \overline{L}_1, Type_1), \ldots, (L_n, \overline{L}_n, Type_n)$ be a pseudo-model for φ and let f be an $\mathsf{L}\overline{\mathsf{L}}$-witness such that $|f(i,j)| > 2 \cdot |\varphi|$ for some pair of indexes i, j. We show how to find a new witness f' such that for each pair of indexes i, j, with $i \leq j$, $|f(i,j)| \leq 2 \cdot |\varphi|$.

Let $i \leq j$ be such that $|f(i,j)| > 2 \cdot |\varphi|$. To fix such a defect, we replace $f(i,j)$ by a subset of it. In doing that, we must be careful not to spoil featuring and fulfillingness conditions of f (notice that, by replacing $f(i,j)$ with a subset of it, there is no danger of spoiling consistency). Let $R_{ij} = \{\psi \in \mathcal{Cl}(\varphi) \mid \exists F \in f(i,j) \text{ with } \psi \in F\}$ be the set of formulae in $\mathcal{Cl}(\varphi)$ that are fulfilled in $f(i,j)$. It clearly holds that $|R_{ij}| \leq |\mathcal{Cl}(\varphi)| \leq 2 \cdot |\varphi|$. We take a minimal subset S of $f(i,j)$ such that, for each $\psi \in R_{ij}$, $\psi \in F$ for some $F \in S$. Since $|S| \leq |R_{ij}| \leq 2 \cdot |\varphi|$, a new $\mathsf{L\overline{L}}$-witness f', with $|f'(i,j)| \leq 2 \cdot |\varphi|$, can be obtained from f as follows:

$$f'(i',j') = \begin{cases} S & \text{if } i = i' \text{ and } j = j'; \\ f(i',j') & \text{otherwise.} \end{cases}$$

By repeating the same process for each pair i', j' such that $|f(i',j')| > 2 \cdot |\varphi|$, we obtain an $\mathsf{L\overline{L}}$-witness f' with $|f'(i,j)| \leq 2 \cdot |\varphi|$ for all $1 \leq i \leq j \leq n$. $\quad\square$

Lemma 2 allows us to to restrict our attention to $\mathsf{L\overline{L}}$-witnesses with polynomially bounded images, that is, we can focus on candidate $\mathsf{L\overline{L}}$-witnesses f consisting of a sequence of at most $\frac{(8 \cdot |\varphi| - 1) \cdot 8 \cdot |\varphi|}{2}$ sets of atoms, each of them of cardinality at most $2 \cdot |\varphi|$. An NP procedure to decide the satisfiability of an $\mathsf{L\overline{L}}$ formula can be easily obtained as follows: we guess an $\mathsf{L\overline{L}}$-sequence of size at most $8 \cdot |\varphi| - 1$ and an $\mathsf{L\overline{L}}$-witness f of size at most $\frac{(8 \cdot |\varphi| - 1) \cdot 16 \cdot |\varphi|^2}{2}$ and then we check whether it fulfills cardinality, featuring, consistency, and fulfillingness conditions.

Theorem 2. *The satisfiability problem for $\mathsf{L\overline{L}}$ over $\mathbb{Q}$ is NP-complete.*

5.3. The fragment $\mathsf{B\overline{B}L\overline{L}}$

In this section, we generalize the proof for $\mathsf{L\overline{L}}$ to $\mathsf{B\overline{B}L\overline{L}}$. In particular, in Lemma 1, we showed how to partition $\mathbb{Q}$ into a finite set of clusters, each one consisting of elements with the same set of $\mathsf{L\overline{L}}$-requests, and then to obtain a pseudo-model from such a partition. The same construction can be exploited in the case of $\mathsf{B\overline{B}L\overline{L}}$, provided that additional information is associated with the clusters of the partition, or, equivalently, the triples of the $\mathsf{L\overline{L}}$-sequence of the pseudo-model.

Let φ be a $\mathsf{B\overline{B}L\overline{L}}$ formula, M be a (concrete) model for it, and x be a rational number. For every interval $[x,y]$, we define the $\mathsf{B\overline{B}}$-*requests* of $[x,y]$ as the pair of sets $(B_{x,y}, \overline{B}_{x,y})$, respectively called B-*requests* and $\overline{\mathsf{B}}$-*requests*, where $B_{x,y}$ (resp., $\overline{B}_{x,y}$) contains all formulae $\psi \in \mathcal{Cl}(\varphi)$ such that

$\langle B \rangle \psi$ (resp., $\langle \overline{B} \rangle \psi$) is true over $[x, y]$. By the transitivity of $\langle B \rangle$ (resp., $\langle \overline{B} \rangle$), for any given left endpoint x, the set of B-requests (resp., $\overline{B}$-requests) is monotone with respect to the ordering of the right endpoints, that is, for every pair y, y', with $y < y'$, it holds that $B_{x,y} \subseteq B_{x,y'}$ (resp., $\overline{B}_{x,y} \supseteq \overline{B}_{x,y'}$). Once again, since the closure of a formula is a finite set, we can partition the set of all intervals with a given left endpoint x into a finite number of clusters of intervals whose elements have the same $B\overline{B}$-requests ($B\overline{B}$-clusters). The partition can be defined in such a way that each cluster is either a singleton or a convex, open set of intervals.

Definition 3. *A $B\overline{B}$-sequence is a finite sequence of quadruples of the form:*

$$\sigma_{B\overline{B}} = (B_1, \overline{B}_1, Cluster_1, BType_1), \ldots, (B_m, \overline{B}_m, Cluster_m, BType_m),$$

where, for $1 \leq i \leq m$, B_i and $\overline{B}_i$ are sets of B-requests and $\overline{B}$-requests, respectively, $Cluster_i \geq 1$, and $BType_i \in \{open, singleton\}$, such that:
   (i) $B_i \subseteq B_{i+1}$ and $\overline{B}_i \supseteq \overline{B}_{i+1}$, for $1 \leq i < m$ (monotonicity of $B\overline{B}$-requests);
   (ii) for each $1 \leq i < m$, either $Cluster_{i+1} = Cluster_i$ or $Cluster_{i+1} = Cluster_i + 1$ (monotonicity of the sequence $Cluster_1, \ldots, Cluster_m$);
   (iii) for all i, j, with $1 \leq i < j \leq m$, if $BType_i = BType_j = open$, then $(B_i, \overline{B}_i) \neq (B_j, \overline{B}_j)$ or $Cluster_i \neq Cluster_j$ (maximality);
   (iv) $BType_1 = BType_m = open$ (the first and last clusters are open sets);
   (v) for each $1 < i < m$, if $BType_i = singleton$, then $BType_{i-1} = BType_{i+1} = open$ (no consecutive clusters of type singleton).

A $B\overline{B}$-sequence models the behaviour of all intervals with the same left endpoint and it is not an independent element of the pseudo-model: it must be embedded into an $L\overline{L}$-sequence in a way that guarantees the consistency of both $B\overline{B}$- and $L\overline{L}$-requests.

Let $\sigma_{B\overline{B}}$ be the $B\overline{B}$-sequence for the intervals with left endpoint x. We call the $L\overline{L}$-cluster $(L_i, \overline{L}_i, Type_i)$ to which x belongs the *starting cluster* of $\sigma_{B\overline{B}}$, and we constrain all the right endpoints included in a $B\overline{B}$-cluster of $\sigma_{B\overline{B}}$ to belong to the same $L\overline{L}$-cluster. $B\overline{B}$-sequences can thus be viewed as a "refinement" of $L\overline{L}$ ones. A graphical account of the resulting picture is given in Figure 4.

The interplay between $B\overline{B}$- and $L\overline{L}$-sequences can be formalized as follows.

Definition 4. *Given an $L\overline{L}$-sequence $\sigma_{L\overline{L}} = (L_1, \overline{L}_1, Type_1), \ldots, (L_n, \overline{L}_n, Type_n)$ and a $B\overline{B}$-sequence $\sigma_{B\overline{B}} = (B_1, \overline{B}_1, Cluster_1, BType_1), \ldots, (B_m, \overline{B}_m,$*
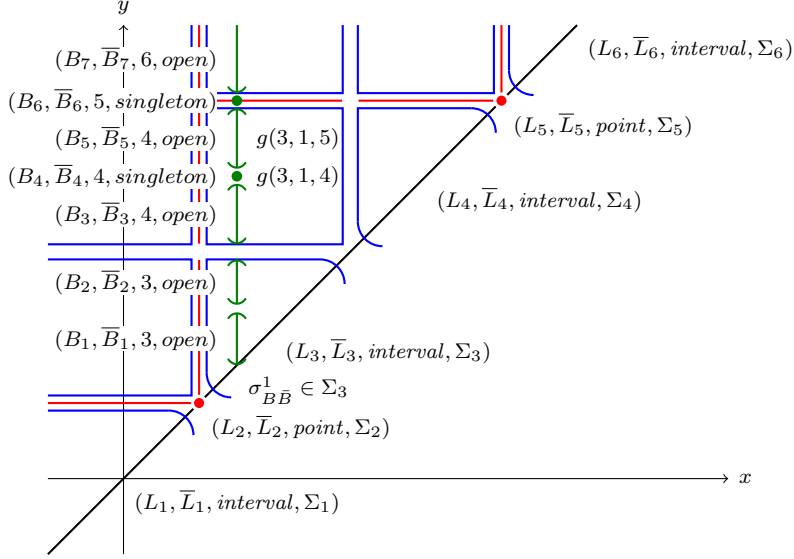
Figure 4: The interplay between $B\overline{B}$- and $L\overline{L}$-sequences: a $B\overline{B}L\overline{L}$ pseudo-model.

$Cluster_m, BType_m)$, we say that $\sigma_{B\bar{B}}$ is compatible with the $L\overline{L}$-cluster $(L_i, \overline{L}_i, Type_i)$ if (and only if):

- $Cluster_1 = i$ (the $B\overline{B}$-sequence starts at $(L_i, \overline{L}_i, Type_i)$);
- $Cluster_m = n$ (the $B\overline{B}$-sequence covers the entire suffix $(L_i, \overline{L}_i, Type_i)$, $\dots, (L_n, \overline{L}_n, Type_n)$ of $\sigma_{L\bar{L}}$);
- if $Type_{Cluster_1} = $ interval, then $BType_1 = $ open (matching types between clusters - condition 1);
- for each $1 \leq j \leq m$, if $Type_{Cluster_j} = $ point, then $BType_j = $ singleton (matching types between clusters - condition 2);
- for each $1 < j < m$, if $Type_{Cluster_j} = $ interval and $Cluster_{j-1} \neq Cluster_j$ or $Cluster_{j+1} \neq Cluster_j$, then $BType_j = $ open (matching types between clusters - condition 3);
- $BType_m = $ open (matching types between clusters - condition 4).

It can be easily checked that there are at most $24 \cdot |\varphi| - 3$ distinct $B\overline{B}$-clusters in the sequence $\sigma_{B\bar{B}}$: by items (i), (ii), and (iii) of Definition 3, there are at most $4 \cdot |\varphi| + (8 \cdot |\varphi| - 1) = 12 \cdot |\varphi| - 1$ clusters of type *open* (notice that for each i, $Cluster_i \leq n \leq 8 \cdot |\varphi| - 1$), and, by items (iv) and (v) of Definition 3, there are at most $12 \cdot |\varphi| - 2$ clusters of type *singleton*. A $B\overline{B}L\overline{L}$-sequence is obtained from an $L\overline{L}$ one by adding a set of compatible $B\overline{B}$-sequences to

each of its clusters.

Definition 5. *Let* $(L_1, \overline{L}_1, Type_1), \ldots, (L_n, \overline{L}_n, Type_n)$ *be an* $\mathsf{L\overline{L}}$*-sequence. A* $\mathsf{B\overline{B}L\overline{L}}$*-sequence is an n-long sequence of quadruples of the form:*

$$\sigma_{\mathsf{B\overline{B}L\overline{L}}} = (L_1, \overline{L}_1, Type_1, \Sigma_1), \ldots, (L_n, \overline{L}_n, Type_n, \Sigma_n)$$

where, for each $1 \le i \le n$, Σ_i *is a finite enumeration of* $\mathsf{B\overline{B}}$*-sequences* $\{\sigma_{B\bar{B}}^{i,1}, \ldots, \sigma_{B\bar{B}}^{i,p_i}\}$ *such that (i) for each* $1 \le j \le p_i$, $\sigma_{B\bar{B}}^{i,j}$ *is compatibile with the cluster* $(L_i, \overline{L}_i, Type_i)$ *and (ii) if* $Type_i = $ point, *then* $p_i = 1$, *otherwise* $p_i \ge 1$.

A $\mathsf{B\overline{B}L\overline{L}}$ *pseudo-model* pairs a $\mathsf{B\overline{B}L\overline{L}}$-sequence with a labeling function.

Definition 6. *A pseudo-model for a* $\mathsf{B\overline{B}L\overline{L}}$ *formula* φ *is a labeled* $\mathsf{B\overline{B}L\overline{L}}$*-sequence* $(L_1, \overline{L}_1, Type_1, \Sigma_1), \ldots, (L_n, \overline{L}_n, Type_n, \Sigma_n)$, *where the labeling function* $g : \{1, \ldots, n\} \times \{1, \ldots, P\} \times \{1, \ldots, M\} \mapsto 2^{atoms(\varphi)}$, *with* $P = \max\{p_i \mid 1 \le i \le n\}$ *and* $M = \max\{|\sigma_{B\bar{B}}| \mid \sigma_{B\bar{B}} \in \Sigma_i, 1 \le i \le n\}$, *satisfies the following constraints on the cardinality of the images:*

- $|g(i,j,k)| = 0$ *if* $|\Sigma_i| < j$ *or* $|\sigma_{B\bar{B}}^{i,j}| < k$ *(incorrect indexes);*
- $|g(i,j,k)| = 1$ *if the k-th* $\mathsf{B\overline{B}}$*-cluster of* $\sigma_{B\bar{B}}^{i,j}$ *is of type* singleton*;*
- $|g(i,j,k)| \ge 1$ *if the k-th* $\mathsf{B\overline{B}}$*-cluster of* $\sigma_{B\bar{B}}^{i,j}$ *is of type* open*,*

as well as the following featuring, consistency, and fulfillingness conditions (for each triple (i,j,k), *we shall denote by* $(B_k, \overline{B}_k, Cluster_k, BType_k)$ *the k-th cluster of* $\sigma_{B\bar{B}}^{i,j}$*):*

- *there exist a triple* (i,j,k) *and an atom* $F \in g(i,j,k)$ *such that* $\varphi \in F$ *(*featuring condition*);*
- *for each triple* (i,j,k) *and atom* $F \in g(i,j,k)$, $\mathcal{R}eq_{\overline{L}}(F) = \overline{L}_i$, $\mathcal{R}eq_L(F) = L_{Cluster_k}$, *if the* $Cluster_k$*-th* $\mathsf{L\overline{L}}$*-cluster is of type* point *(resp., interval), then* $\mathcal{O}bs(F) \subseteq \overline{L}_{Cluster_k+1}$ *(resp.,* $\mathcal{O}bs(F) \subseteq \overline{L}_{Cluster_k}$*), and if the i-th* $\mathsf{L\overline{L}}$*-cluster is of type* point *(resp., interval), then* $\mathcal{O}bs(F) \subseteq L_{i-1}$ *(resp.,* $\mathcal{O}bs(F) \subseteq L_i$*) (*$\mathsf{L\overline{L}}$*-consistency);*
- *for each triple* (i,j,k) *and atom* $F \in g(i,j,k)$, $\mathcal{R}eq_B(F) = B_k$, $\mathcal{R}eq_{\overline{B}}(F) = \overline{B}_k$, *and if* $BType_k = $ singleton *(resp., open), then* $\mathcal{O}bs(F) \subseteq \overline{B}_{k-1} \cap B_{k+1}$ *(resp.,* $\mathcal{O}bs(F) \subseteq \overline{B}_k \cap B_k$*) (*$\mathsf{B\overline{B}}$*-consistency);*
- *for each index i, with* $1 \le i \le n$, *and each formula* $\psi \in L_i$, *if the i-th* $\mathsf{L\overline{L}}$*-cluster is of type* point *(resp., interval), there exist a triple* (i', j', k'), *with* $i < i'$ *(resp.,* $i \le i'$*), and an atom* $F \in g(i', j', k')$ *such that* $\psi \in F$ *(*L*-fulfillingness);*

- *for each index i, with $1 \leq i \leq n$, and each formula $\psi \in \overline{L}_i$, if the i-th $\mathsf{L}\overline{\mathsf{L}}$-cluster is of type* point *(resp.,* interval*), there exist a triple (i', j', k'), with $i' < i$ and $Cluster_{k'} < i$ (resp., $i' \leq i$ and $Cluster_{k'} \leq i$), and an atom $F \in g(i', j', k')$ such that $\psi \in F$ ($\overline{\mathsf{L}}$-fulfillingness);*
- *for each triple (i, j, k), with $1 \leq i \leq n$, $\sigma^{i,j}_{B\bar{B}} \in \Sigma_i$, and $|\sigma^{i,j}_{B\bar{B}}| \geq k$, and formula $\psi \in B_k$, if the k-th $\mathsf{B}\overline{\mathsf{B}}$-cluster is of type* singleton *(resp.,* open*), there exist an index k', with $k' < k$ (resp., $k' \leq k$), and an atom $F \in g(i, j, k')$ such that $\psi \in F$ (B-fulfillingness);*
- $\overline{\mathsf{B}}$*-fulfillingness is defined analogously.*

We call the labeling function g a $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$-*witness* for the $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$-sequence $(L_1, \overline{L}_1, Type_1, \Sigma_1), \ldots, (L_n, \overline{L}_n, Type_n, \Sigma_n)$. Figure 4 gives an intuitive account of how to expand the $\mathsf{L}\overline{\mathsf{L}}$ pseudo-model with 6 $\mathsf{L}\overline{\mathsf{L}}$-clusters of Figure 3 into a full $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$ pseudo-model. In particular, it displays a $\mathsf{B}\overline{\mathsf{B}}$-sequence (in green) that is compatible with the 3rd $\mathsf{L}\overline{\mathsf{L}}$-cluster, which partitions the half-plane in regions that can be either "open segments", e.g., $g(3, 1, 4)$, or "points", e.g., $g(3, 1, 5)$. Cardinality, featuring, consistency, and fulfillingness conditions guarantees that the labeling defined by the function g respects the semantics of the formulae and the shape of the different regions.

Lemma 3. *A $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$ formula φ is satisfiable over $\mathbb{Q}$ if and only if there exists a $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$ pseudo-model for it.*

Proof. We first prove the left-to-right implication. Let $\mathbf{L} = (\mathbb{Q}, \mathcal{L})$ be a fulfilling *LIS* for φ, whose existence is guaranteed by Theorem 1. We build a $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$ pseudo-model for φ as follows. Let $[\![x_1]\!]$, $[\![x_2]\!]$, $\ldots$, $[\![x_n]\!]$ be the partition of $\mathbb{Q}$ and $\sigma_{L\bar{L}} = (L_1, \overline{L}_1, Type_1), (L_2, \overline{L}_2, Type_2), \ldots, (L_n, \overline{L}_n, Type_n)$ be the corresponding $\mathsf{L}\overline{\mathsf{L}}$-sequence defined as in the proof of Lemma 1.

For each $x \in \mathbb{Q}$, we define a finite partition $\langle\!\langle x, y_1 \rangle\!\rangle, \langle\!\langle x, y_2 \rangle\!\rangle, \ldots, \langle\!\langle x, y_m \rangle\!\rangle$ of the intervals with left endpoint x that satisfies the following properties (for each i, we choose $[x, y_i] \in \langle\!\langle x, y_i \rangle\!\rangle$ as the representative of the class $\langle\!\langle x, y_i \rangle\!\rangle$):

- for all i, $\langle\!\langle x, y_i \rangle\!\rangle$ is either the singleton $\{[x, y_i]\}$ or an open set of intervals $\{[x, y] \mid y \in (y_i^-, y_i^+)$ for some open interval $(y_i^-, y_i^+)\}$ (notice that y_i^+ might be $+\infty$);
- for all i, all intervals in $\langle\!\langle x, y_i \rangle\!\rangle$ have the same set of $\mathsf{B}\overline{\mathsf{B}}$-requests $(B_i, \overline{B}_i)$; moreover, if $y_i \in [\![x_j]\!]$, so does $\bar{y}$, for all $[x, \bar{y}] \in \langle\!\langle x, y_i \rangle\!\rangle$;
- for all pairs of open sets $\langle\!\langle x, y_i \rangle\!\rangle, \langle\!\langle x, y_j \rangle\!\rangle$, with $i \neq j$, $(B_i, \overline{B}_i) \neq (B_j, \overline{B}_j)$ or $y_i \in [\![x_{i'}]\!]$ and $y_j \in [\![x_{j'}]\!]$ for some $i' \neq j'$;

- the set of witnesses of the partition is monotonically increasing: $y_1 < y_2 < \ldots < y_m$.

For all $x \in \mathbb{Q}$, the above partition induces a $B\overline{B}$-sequence:

$$\sigma_{B\overline{B}}^x = (B_1, \overline{B}_1, Cluster_1, BType_1), \ldots, (B_m, \overline{B}_m, Cluster_m, BType_m)$$

where $Cluster_i$ is such that $y_i \in [\![x_{Cluster_i}]\!]$ and $BType_i = singleton$ if $\langle\!\langle x, y_i \rangle\!\rangle$ is a singleton, $BType_i = open$ otherwise. It can be easily shown that if $x \in [\![x_i]\!]$, then $\sigma_{B\overline{B}}^x$ is compatible with the i-th $L\overline{L}$-cluster $(L_i, \overline{L}_i, Type_i)$. Hence, for each cluster $(L_i, \overline{L}_i, Type_i)$ of the $L\overline{L}$-sequence $\sigma_{L\overline{L}}$ we can define the set of induced $B\overline{B}$-sequences $\Sigma_i = \{\sigma_{B\overline{B}}^x \mid x \in [\![x_i]\!]\}$. Even though $[\![x_i]\!]$ may contain infinitely many rational numbers, there are only finitely many different $B\overline{B}$-sequences compatible with $(L_i, \overline{L}_i, Type_i)$, and thus Σ_i is a finite set that we can enumerate as $\Sigma_i = \{\sigma_{B\overline{B}}^{i,1}, \ldots, \sigma_{B\overline{B}}^{i,p_i}\}$.

We define the $B\overline{B}L\overline{L}$-sequence of the $B\overline{B}L\overline{L}$ pseudo-model for φ as follows:

$$\sigma_{\mathsf{B\overline{B}L\overline{L}}} = (L_1, \overline{L}_1, Type_1, \Sigma_1), (L_2, \overline{L}_2, Type_2, \Sigma_2), \ldots, (L_n, \overline{L}_n, Type_n, \Sigma_n),$$

while the labeling function g (the $B\overline{B}L\overline{L}$-witness) is defined as follows:

$$g(i, j, k) = \{\mathcal{L}([x, y]) \mid \sigma_{B\overline{B}}^{i,j} \in \Sigma_i, \sigma_{B\overline{B}}^x = \sigma_{B\overline{B}}^{i,j} \text{ and } [x, y] \in \langle\!\langle x, y_k \rangle\!\rangle\},$$

if $|\Sigma_i| \geq j$ and $|\sigma_{B\overline{B}}^{i,j}| \geq k$; $g(i, j, k) = \emptyset$ otherwise. It is easy to check that g fulfills the cardinality, featuring, consistency, and fulfillingness constraints.

To prove the opposite implication, let us consider a $B\overline{B}L\overline{L}$ pseudo-model for φ consisting of the $B\overline{B}L\overline{L}$-sequence $(L_1\overline{L}_1, Type_1, \Sigma_1), \ldots, (L_n, \overline{L}_n, Type_n, \Sigma_n)$ and the associated labeling function g, that satisfies the cardinality, featuring, consistency, and fulfillingness conditions. As in the proof of Lemma 1, the fulfilling LIS $\mathbf{L} = (\mathbb{Q}, \mathcal{L})$ satisfying φ is obtained as the limit of an infinite sequence of finite LISs $\mathbf{L}_0 = (\mathbb{D}_0, \mathcal{L}_0), \mathbf{L}_1 = (\mathbb{D}_1, \mathcal{L}_1), \mathbf{L}_2 = (\mathbb{D}_2, \mathcal{L}_2), \ldots$. We pair each finite LIS $\mathbf{L}_i = (\mathbb{D}_i, \mathcal{L}_i)$ with (i) a function $Clst_i$ that associates elements in D_i with $L\overline{L}$-clusters (as in Lemma 1), (ii) a function $Seq_i : D_i \mapsto \mathbb{N}$ that associates each $x \in \mathbb{D}_i$ with a $B\overline{B}$-sequence $\sigma_{B\overline{B}}^{Clst_i(x), Seq_i(x)} \in \Sigma_{Clst_i(x)}$, and (iii) a function $IClst_i : \mathbb{I}(\mathbb{D}_i) \mapsto \mathbb{N}$ that associates each interval $[x, y] \in \mathbb{I}(\mathbb{D}_i)$ with a $B\overline{B}$-cluster in the $B\overline{B}$-sequence $\sigma_{B\overline{B}}^{Clst_i(x), Seq_i(x)}$. The three functions are such that for each interval $[x, y] \in \mathbb{I}(\mathbb{D}_i)$, $\mathcal{L}_i([x, y]) \in g(Clst_i(x), Seq_i(x), IClst_i([x, y]))$.

The initial *LIS* $\mathbf{L}_0 = (\mathbb{D}_0, \mathcal{L}_0)$ is such that the domain $D_0 = \{1, 2, \ldots, n\}$ contains an element for each L$\overline{\text{L}}$-cluster and the labeling function $\mathcal{L}_0$ features the formula φ (its definition is similar to the one given in the proof of Lemma 1, and thus we omit its details). Then, the *LIS* $\mathbf{L}_{i+1}$ is obtained from $\mathbf{L}_i$ by choosing one of its defects and fixing it.

L-defects, $\overline{\text{L}}$-defects, Disc-defects, Min-defects, and Max-defects are dealt with as in the proof of Lemma 1. In addition, we must cope with B- and $\overline{\text{B}}$-requests which are not fulfilled (B-defects and $\overline{\text{B}}$-defects, respectively). These defects are fixed as follows.

B-defects: a B-defect occurs if there exist an interval $[x_a, x_b] \in \mathbb{I}(\mathbb{D}_i)$ and a formula $\langle B \rangle \psi \in \mathcal{L}_i([x_a, x_b])$ such that for all $[x_a, x_c] \in \mathbb{I}(\mathbb{D}_i)$, with $x_c < x_b$, $\psi \notin \mathcal{L}_i([x_a, x_c])$.

To fix such a defect, a new interval that satisfies ψ must be properly added to $\mathbf{L}_{i+1}$. Let $i' = Clst_i(x_a)$, $j' = Seq_i(x_a)$, and $k' = IClst_i([x_a, x_b])$. Moreover, for each k, let $(B_k, \overline{B}_k, Cluster_k, BType_k)$ be the k-th cluster of the j'-th B$\overline{\text{B}}$-sequence $\sigma_{B\overline{B}}^{i',j'} \in \Sigma_{i'}$. We have that $\mathcal{L}_i([x_a, x_b]) \in g(i', j', k')$. By the B$\overline{\text{B}}$-consistency of g, $\mathcal{R}eq_B(\mathcal{L}_i([x_a, x_b])) = B_{k'}$, which implies that $\psi \in B_{k'}$. By the B-fulfillingness condition, there exist an index k'' and an atom $F_\psi \in g(i', j', k'')$ such that $k'' \le k'$ and $\psi \in F_\psi$. Let x_c be the greatest element in $\mathbb{D}_i$ such that $IClst([x_a, x_c]) = k''$ and let x'_c be its immediate successor in $\mathbb{D}_i$. The right endpoint of the new interval $[x_a, y']$ is defined as follows:

$$y' = \begin{cases} x_c + 1 & \text{if } x_c = \max(\mathbb{D}_i); \\ \frac{x_c + x'_c}{2} & \text{otherwise.} \end{cases}$$

Hence, we let $D_{i+1} = D_i \cup \{y'\}$. To complete the definition of $\mathbf{L}_{i+1}$, we put $i'' = Cluster_{k''}$ and we proceed as follows: (i) we set $Clst_{i+1}(y') = i''$ (y' belongs to the i''-th L$\overline{\text{L}}$-cluster); (ii) we associate a (randomly selected) B$\overline{\text{B}}$-sequence $\sigma_{B\overline{B}}^{i'',j''}$ from $\Sigma_{i''}$ with y' by setting $Seq_{i+1}(y') = j''$ for some $1 \le j'' \le |\Sigma_{i''}|$; (iii) we set $IClst([y', x]) = k'''$, where the k'''-th B$\overline{\text{B}}$-cluster $(B_{k'''}, \overline{B}_{k'''}, Cluster_{k'''}, BType_{k'''})$ of the j''-th B$\overline{\text{B}}$-sequence $\sigma_{B\overline{B}}^{i'',j''} \in \Sigma_{i''}$ is such that $Cluster_{k'''} = Clst_i(x)$; and (iv) we set $IClst([x, y']) = k'''$, where the k'''-th B$\overline{\text{B}}$-cluster $(B_{k'''}, \overline{B}_{k'''}, Cluster_{k'''}, BType_{k'''})$ of the $Seq_i(x)$-th B$\overline{\text{B}}$-sequence $\sigma_{B\overline{B}}^{Clst_i(x),Seq_i(x)} \in \Sigma_{Clst_i(x)}$ is such that $Cluster_{k'''} = i''$. As for the labeling $\mathcal{L}_{i+1}$, we let $\mathcal{L}_{i+1}([x_a, y']) = F_\psi$ and for all new intervals $[x, y]$, that is, intervals $[x, y]$, with $x = y'$ or $y = y'$, we put $\mathcal{L}_{i+1}([x, y]) \in g(Clst_{i+1}(x), Seq_{i+1}(x), IClst_{i+1}([x, y]))$; the labeling of the other intervals does not change.

$\overline{\mathsf{B}}$-**defects** are dealt with in a completely symmetrical way, and thus details are omitted.

By fixing all the defects according to a fair strategy which guarantees that no $\langle L\rangle$-/$\langle\overline{L}\rangle$-/$\langle B\rangle$-/$\langle\overline{B}\rangle$-request is indefinitely postponed and the final domain is dense, we can force the limit structure $\mathbf{L}_\omega = \bigcup_{i\in\mathbb{N}} \mathbf{L}_i$ to be a fulfilling LIS for φ. $\qquad\square$

To prove that the satisfiability problem for $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$ is in NP, we need to show that the existence of a $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$ pseudo-model can be checked in polynomial time. We have already shown that the length of an $\mathsf{L}\overline{\mathsf{L}}$-sequence and of a $\mathsf{B}\overline{\mathsf{B}}$-sequence are bounded by $8\cdot|\varphi|-1$ and $24\cdot|\varphi|-3$, respectively. However, this is not enough. Polynomial bounds on the number of $\mathsf{B}\overline{\mathsf{B}}$-sequences contained in the Σ_i component of an $\mathsf{L}\overline{\mathsf{L}}$-cluster and on the size of the $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$ witness function g are needed. The next lemma provides such missing bounds.

Lemma 4. *Let φ be a $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$ formula which is satisfiable over $\mathbb{Q}$. Then, there exists a $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$ pseudo-model for φ consisting of a $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$-sequence $(L_1, \overline{L}_1, Type_1, \Sigma_1), \ldots, (L_n, \overline{L}_n, Type_n, \Sigma_n)$ and an associated $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$-witness g such that (i) $|\Sigma_i| \leq 2\cdot|\varphi|$, for each $1\leq i\leq n$, and (ii) $|g(i,j,k)| \leq 2\cdot|\varphi|$, for each triple (i,j,k) in the domain of g.*

Proof. Let us prove item (i). Let $(L_1, \overline{L}_1, Type_1, \Sigma_1), \ldots, (L_n, \overline{L}_n, Type_n, \Sigma_n)$ be a $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$ pseudo-model that violates it, that is, $|\Sigma_i| > 2\cdot|\varphi|$, for some $1\leq i\leq n$. We proceed as in the proof of Lemma 2. Let $R_i = \{\psi \in \mathcal{C}l(\varphi) \mid \exists j,k, \text{ and } F \text{ such that } F \in g(i,j,k) \text{ and } \psi \in F\}$. It holds that $|R_i| \leq 2\cdot|\varphi|$. Let S_i be a minimal subset of Σ_i such that, for each $\psi \in R_i$, there is $\sigma_{B\overline{B}}^{i,j} \in S_i$ that fulfills ψ, that is, $\psi \in g(i,j,k))$ for some k. Since $|S_i| \leq 2\cdot|\varphi|$, we can replace Σ_i by S_i in the pseudo-model. By repeating the same process to all i such that $|\Sigma_i| > 2\cdot|\varphi|$, we obtain a pseudo-model that fulfills condition (i). Item (ii) can be proved in a similar way, by replacing each $g(i,j,k)$ such that $|g(i,j,k)| > 2\cdot|\varphi|$ with a minimal subset fulfilling all $\mathsf{B}\overline{\mathsf{B}}$-requests. $\qquad\square$

Thanks to Lemma 4, we can restrict our attention to $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$ pseudo-models of polynomially-bounded size. First, the number of quadruples of the $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$-sequence of a $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$ pseudo-model is equal to the number of clusters of the $\mathsf{L}\overline{\mathsf{L}}$-sequence and thus it is bounded by $8\cdot|\varphi|-1$. Then, by Lemma 4 (item (i)), for each quadruple $(L_i, \overline{L}_i, Type_i, \Sigma_i)$ in the $\mathsf{B}\overline{\mathsf{B}}\mathsf{L}\overline{\mathsf{L}}$-sequence, we can assume the number of $\mathsf{B}\overline{\mathsf{B}}$-sequences in Σ_i to be bounded by $2\cdot|\varphi|$. Finally, the number of clusters in each $\sigma_{B\overline{B}}$-sequence is bounded by $24\cdot|\varphi|-3$. Hence,

the total number of $\sigma_{B\bar{B}}$-clusters in the $\mathsf{B\overline{B}L\overline{L}}$ pseudo-model, or, equivalently, the total number of meaningful triples (i, j, k), is bounded by $2 \cdot |\varphi| \cdot (8 \cdot |\varphi| - 1) \cdot (24 \cdot |\varphi| - 3) = 384 \cdot |\varphi|^3 - 96 \cdot |\varphi|^2 + 6 \cdot |\varphi|$. Now, by Lemma 4 (item (ii)), for each triple (i, j, k), $|g(i, j, k)| \leq 2 \cdot |\varphi|$ and thus the size of g is at most $2 \cdot |\varphi| \cdot (384 \cdot |\varphi|^3 - 96 \cdot |\varphi|^2 + 6 \cdot |\varphi|)$. An NP procedure to decide the satisfiability of a $\mathsf{B\overline{B}L\overline{L}}$ formula can be obtained as follows: we first guess a $\mathsf{B\overline{B}L\overline{L}}$-sequence and an associated $\mathsf{B\overline{B}L\overline{L}}$-witness, that satisfy the above bounds, and then we check whether cardinality, featuring, consistency, and fulfillingness conditions are fulfilled.

Theorem 3. *The satisfiability problem for $\mathsf{B\overline{B}L\overline{L}}$ over $\mathbb{Q}$ is NP-complete.*

6. EXPSPACE-complete fragments

In this section, we show that the satisfiability problem for AB (and $\mathsf{A\overline{B}}$) over $\mathbb{Q}$ is EXPSPACE-hard via a reduction from the *exponential-corridor tiling problem*, which is known to be EXPSPACE-complete [28].

Formally, an instance of the exponential-corridor tiling problem is a tuple $\mathcal{T} = (T, t_\perp, t_\top, H, V, n)$, where T is a finite set of tiles, $t_\perp \in T$ is the bottom tile, $t_\top \in T$ is the top tile, H and V are two binary relations over T, that specify the horizontal and vertical constraints, respectively, and n is a positive natural number encoded in unary notation. The problem consists of establishing whether there is a tiling function $f : \mathbb{N} \times \{0, \ldots, 2^n - 1\} \mapsto T$ of the infinite discrete corridor of height 2^n such that: *(i)* for all $x \in \mathbb{N}$, $f(x, 0) = t_\perp$ (it associates the tile $t_\perp$ with the bottom row of the corridor); *(ii)* for all $x \in \mathbb{N}$, $f(x, 2^n - 1) = t_\top$ (it associates the tile $t_\top$ with the top row of the corridor); *(iii)* for all $x \in \mathbb{N}$ and $0 \leq y < 2^n$, $(f(x, y), f(x + 1, y)) \in H$ (it satisfies the horizontal constraints H); *(iv)* for all $x \in \mathbb{N}$ and $0 \leq y < 2^n - 1$, $(f(x, y), f(x, y + 1)) \in V$ (it satisfies the vertical constraints V).

In the following, we reduce the exponential-corridor tiling problem to the satisfiability problem for AB. The reduction exploits: *(a)* a correspondence between the points $p = (x, y)$ of the infinite corridor $\mathbb{N} \times \{0, ..., 2^n - 1\}$ and a suitable set of "unit" intervals labeled by the proposition letter u; *(b)* n additional proposition letters to encode the binary representation of the y-coordinate of each row of the corridor; *(c)* $|T|$ proposition letters to represent tiles in T; *(d)* modalities $\langle A \rangle$ and $\langle B \rangle$ to enforce the constraints on f.

For an interval $[x, y]$, let $\mathcal{G}^{\mathsf{AB}}_{[x,y]}$ be the set of intervals $\{[x, y]\} \cup \{[w, z] \mid w \geq y\}$. In the rest of the section, even when not explicitly said, we focus

on intervals in $\mathcal{G}^{\mathsf{AB}}_{[x,y]}$. Let $[G^{\mathsf{AB}}]\varphi$ be the following shortcut: $[G^{\mathsf{AB}}]\varphi = \varphi \wedge [A]\varphi \wedge [A][A]\varphi$. Modality $[G^{\mathsf{AB}}]$ can be thought of as the global modality for the fragment AB as it allows one to predicate about all intervals in the relevant set $\mathcal{G}^{\mathsf{AB}}_{[x,y]}$. This is formally stated by the next proposition, whose straightforward proof is omitted (to keep the notation light, and with an abuse of notation, in the rest of the section we omit the superscript $^{\mathsf{AB}}$ in $[G^{\mathsf{AB}}]$ and denote it simply by $[G]$).

Proposition 1. *Let φ be an AB-formula and $[x,y]$ be an interval. Then, $M, [x,y] \Vdash [G]\varphi$ if and only if $M, [w,z] \Vdash \varphi$, for all $[w,z] \in \mathcal{G}^{\mathsf{AB}}_{[x,y]}$.*

We start by building an infinite chain of unit intervals by means of the following formula:

$$\varphi_{u-chain} = \langle A \rangle u \wedge [G](u \to \langle A \rangle u) \wedge [G](u \to \neg \langle B \rangle u) \wedge [G](u \to \neg \langle B \rangle \langle A \rangle u)$$

Lemma 5. *Let $M, [x,y] \Vdash \varphi_{u-chain}$. Then, M features an infinite ascending sequence of points $x_0 < x_1 < x_2 < \ldots$, such that:*

(i) $x_0 = y$;
(ii) $[x_i, x_{i+1}]$ is a u-interval, for every $i \geq 0$;
(iii) if $[w,z]$ is a u-interval, then either $w < x_0$, or $[w,z] = [x_i, x_{i+1}]$ for some $i \geq 0$, or $x_j < w$ for every $j \geq 0$.

Proof. The first conjunct of $\varphi_{u-chain}$ guarantees the existence of a u-interval, say it $[x_0, x_1]$, with $x_0 = y$. The second conjunct states that every u-interval $[w,z]$ is immediately followed by another one $[w',z']$, with $w' = z$. This is enough to prove *(i)* and *(ii)*. To prove *(iii)*, suppose, by contradiction, that there exists a u-interval $[w,z]$ such that $[w,z] \neq [x_i, x_{i+1}]$ for any i and $x_0 \leq w \leq x_j$ for some j. Three cases may arise. If $w = x_i$ and $z < x_{i+1}$, for some $0 \leq i \leq j$, then $[x_i, x_{i+1}]$ is a u-interval that satisfies $\langle B \rangle u$. If $w = x_i$ and $z > x_{i+1}$, for some $0 \leq i \leq j$, then $[x_i, z]$ is a u-interval that satisfies $\langle B \rangle u$. Finally, if $x_i < w < x_{i+1}$, for some $0 \leq i < j$, then $[x_i, w] \Vdash \langle A \rangle u$ and thus $[x_i, x_{i+1}] \Vdash \langle B \rangle \langle A \rangle u$. The first two cases violate the third conjunct of $\varphi_{u-chain}$, the third one violates the fourth conjunct. The thesis follows. $\square$

In the reduction, u-intervals are used to represent the points (x,y) of the infinite corridor $\mathbb{N} \times \{0, \ldots, 2^n - 1\}$. To correctly encode the vertical and horizontal constraints of the tiling function, we need to identify the y coordinate of the points of the corridor. To this end, we introduce the

proposition letters $y_0, \ldots, y_{n-1}$ to represent the binary expansion of the y coordinate and to "count" from 0 to $2^n - 1$. Let $a_{n-1} \ldots a_0$ be the binary expansion of a number a, with $0 \le a \le 2^n - 1$, and let k be the index such that $a_k = 0$ and $a_i = 1$ for all $i < k$, if such an index exists; otherwise (that is, when all digits are equal to 1), we put $k = n$. The binary expansion $b_{n-1} \ldots b_0$ of $a + 1 \mod 2^n$ can be defined as

$$
b_i = \begin{cases} 0 & \text{if } i < k \\ 1 & \text{if } i = k \\ a_i & \text{otherwise} \end{cases}
$$

To encode the above definition of the successor, we make use of two auxiliary formulae φ_{eq}^i and φ_{inc}^i, which are defined by induction on i as follows:

$$
\varphi_{eq}^i = \begin{cases} \top & \text{if } i = n \\ \Big(\big(y_i \wedge \langle A \rangle (u \wedge y_i) \big) \vee \big(\neg y_i \wedge \langle A \rangle (u \wedge \neg y_i) \big) \Big) \wedge \varphi_{eq}^{i+1} & \text{if } i < n \end{cases}
$$

$$
\varphi_{inc}^i = \begin{cases} \top & \text{if } i = n \\ \big(y_i \wedge \langle A \rangle (u \wedge \neg y_i) \wedge \varphi_{inc}^{i+1} \big) \vee \big(\neg y_i \wedge \langle A \rangle (u \wedge y_i) \wedge \varphi_{eq}^{i+1} \big) & \text{if } i < n \end{cases}
$$

Intuitively, φ_{eq}^i forces two adjacent intervals to assign the same truth value to the proposition letters $y_i, \ldots, y_{n-1}$, while φ_{inc}^i guarantees that the binary number represented by the letters $y_i, \ldots, y_{n-1}$ (ignoring the least significant digits) is incremented by one on the next interval. To formally prove that the above formulae match the intuition, we define an auxiliary function $\mathrm{bin}([x, y], j)$ that, for every interval $[x, y]$ and each index $0 \le j < n$, returns the number encoded by the most significant proposition letters $y_j, \ldots, y_{n-1}$:

$$
\mathrm{bin}([x, y], j) = \sum\nolimits_{i \in \{i \in \mathbb{N} \ | \ j \le i < n \text{ and } M, [x,y] \Vdash y_i\}} 2^{i-j}
$$

The next lemma proves that the formulae φ_{eq}^i and φ_{inc}^i actually encode the successor (modulo 2^n) of a y-coordinate.

Lemma 6. *Let $M, [x, y] \Vdash \varphi_{u-chain}$ and let $x_0, x_1, \ldots$ be the sequence of points of Lemma 5. Then, for every $[x_i, x_j]$ and each $0 \le k < n$, it holds that:*

(i) $[x_i, x_j] \Vdash \varphi_{eq}^k$ *if and only if* $\mathrm{bin}([x_j, x_{j+1}], k) = \mathrm{bin}([x_i, x_j], k)$;

38

(ii) $[x_i, x_j] \Vdash \varphi_{inc}^k$ if and only if $\mathrm{bin}([x_j, x_{j+1}], k) = \mathrm{bin}([x_i, x_j], k) + 1$ mod 2^{n-k}.

Proof. The proof is by (backward) induction on k.

Base case: $k = n - 1$. First, assume that $[x_i, x_j] \Vdash \varphi_{eq}^{n-1}$. Two cases may arise. If $[x_i, x_j] \Vdash y_{n-1}$, then $[x_i, x_j] \Vdash \langle A \rangle(u \wedge y_{n-1})$, and thus, by Lemma 5, $[x_j, x_{j+1}] \Vdash y_{n-1}$. Similarly, if $[x_i, x_j] \Vdash \neg y_{n-1}$, then $[x_j, x_{j+1}] \Vdash \neg y_{n-1}$, and thus $\mathrm{bin}([x_j, x_{j+1}], n - 1) = \mathrm{bin}([x_i, x_j], n - 1)$. As for the converse implication, assume that $\mathrm{bin}([x_j, x_{j+1}], n - 1) = \mathrm{bin}([x_i, x_j], n - 1)$. Then, either y_{n-1} is true over both $[x_i, x_j]$ and $[x_j, x_{j+1}]$ or it is false over both of them. Both scenarios imply that $[x_i, x_j] \Vdash \varphi_{eq}^{n-1}$, and thus (i) is proved. To prove (ii), suppose that $[x_i, x_j] \Vdash \varphi_{inc}^{n-1}$. By an argument similar to the above one, we can prove that $[x_i, x_j] \Vdash y_{n-1}$ if and only if $[x_j, x_{j+1}] \Vdash \neg y_{n-1}$, thus showing that $\mathrm{bin}([x_j, x_{j+1}], n-1) = \mathrm{bin}([x_i, x_j], n-1)+1 \mod 2$. As for the converse, suppose that $\mathrm{bin}([x_j, x_{j+1}], n - 1) = \mathrm{bin}([x_i, x_j], n - 1) + 1 \mod 2$. Then, either y_{n-1} is true over $[x_i, x_j]$ and false over $[x_j, x_{j+1}]$, or vice versa. Both scenarios imply that $[x_i, x_j] \Vdash \varphi_{inc}^{n-1}$.

Inductive step: $k < n-1$. Let us assume (i) and (ii) to hold for $k+1$. To prove (i), suppose that $[x_i, x_j] \Vdash \varphi_{eq}^k$. Two cases may arise. If $[x_i, x_j] \Vdash y_k$, then $[x_i, x_j] \Vdash \langle A \rangle(u \wedge y_k)$, and thus, by Lemma 5, $[x_j, x_{j+1}] \Vdash y_k$. Similarly, if $[x_i, x_j] \Vdash \neg y_k$, then $[x_j, x_{j+1}] \Vdash \neg y_k$. Moreover, $[x_i, x_j] \Vdash \varphi_{eq}^k$ implies $[x_i, x_j] \Vdash \varphi_{eq}^{k+1}$, and, by the inductive hypothesis, $\mathrm{bin}([x_j, x_{j+1}], k + 1) = \mathrm{bin}([x_i, x_j], k + 1)$. Since the truth value of y_k is the same over $[x_i, x_j]$ and $[x_j, x_{j+1}]$, we can conclude that $\mathrm{bin}([x_j, x_{j+1}], k) = \mathrm{bin}([x_i, x_j], k)$. To prove the converse implication, suppose that $\mathrm{bin}([x_j, x_{j+1}], k) = \mathrm{bin}([x_i, x_j], k)$. It immediately follows that the truth value of y_k is the same over $[x_i, x_j]$ and $[x_j, x_{j+1}]$, and that $\mathrm{bin}([x_j, x_{j+1}], k + 1) = \mathrm{bin}([x_i, x_j], k + 1)$. It follows that $[x_i, x_j] \Vdash \big(y_i \wedge \langle A \rangle(u \wedge y_i) \big) \vee \big(\neg y_i \wedge \langle A \rangle(u \wedge \neg y_i) \big)$, and, by the inductive hypothesis, we also have that $[x_i, x_j] \Vdash \varphi_{eq}^{k+1}$. Thus, $[x_i, x_j] \Vdash \varphi_{eq}^k$.

To prove (ii), suppose $[x_i, x_j] \Vdash \varphi_{inc}^k$. Two cases may arise. If $[x_i, x_j] \Vdash y_k$, then $[x_i, x_j] \Vdash \varphi_{inc}^{k+1}$ and $[x_j, x_{j+1}] \Vdash \neg y_k$. By the inductive hypothesis, $\mathrm{bin}([x_j, x_{j+1}], k + 1) = \mathrm{bin}([x_i, x_j], k + 1) + 1 \mod 2^{n-(k+1)}$. Then, by definition of bin, $\mathrm{bin}([x_j, x_{j+1}], k) = 2 \cdot \mathrm{bin}([x_j, x_{j+1}], k + 1) + 0 = 2 \cdot (\mathrm{bin}([x_i, x_j], k+1)+1 \mod 2^{n-(k+1)}) = 2\cdot(\mathrm{bin}([x_i, x_j], k+1)+1) \mod 2^{n-k} = \mathrm{bin}([x_i, x_j], k) + 1 \mod 2^{n-k}$. If $[x_i, x_j] \Vdash \neg y_k$, then $[x_i, x_j] \Vdash \varphi_{eq}^{k+1}$ and $[x_j, x_{j+1}] \Vdash y_k$. By the inductive hypothesis, $\mathrm{bin}([x_j, x_{j+1}], k+1)=\mathrm{bin}([x_i, x_j], k+1)$. Then, $\mathrm{bin}([x_j, x_{j+1}], k) = 2\cdot\mathrm{bin}([x_j, x_{j+1}], k+1)+1 = 2\cdot\mathrm{bin}([x_i, x_j], k+1)+1 = \mathrm{bin}([x_i, x_j], k)+1$. Since y_k is false over $[x_i, x_j]$, $\mathrm{bin}([x_i, x_j], k) < 2^{n-k}-1$,

and thus $\text{bin}([x_i, x_j], k) + 1 = \text{bin}([x_i, x_j], k) + 1 \mod 2^{n-k}$.

As for the converse implication, assume $\text{bin}([x_j, x_{j+1}], k)$ to be equal to $\text{bin}([x_i, x_j], k) + 1 \mod 2^{n-k}$. Two cases may arise. If $[x_i, x_j] \Vdash \neg y_k$, then $[x_j, x_{j+1}] \Vdash y_k$ and $\text{bin}([x_j, x_{j+1}], k+1) = \text{bin}([x_i, x_j], k+1)$. Then, it follows that $[x_i, x_j] \Vdash \langle A \rangle (u \wedge y_k)$. By (i), $[x_i, x_j] \Vdash \varphi_{eq}^{k+1}$, and thus $[x_i, x_j] \Vdash \varphi_{inc}^k$. If $[x_i, x_j] \Vdash y_k$, then $[x_j, x_{j+1}] \Vdash \neg y_k$ and $\text{bin}([x_j, x_{j+1}], k+1) = \text{bin}([x_i, x_j], k+1) + 1 \mod 2^{n-(k+1)}$. It follows that $[x_i, x_j] \Vdash \langle A \rangle (u \wedge \neg y_k)$, and, by the inductive hypothesis, $[x_i, x_j] \Vdash \varphi_{inc}^{k+1}$. Thus, $[x_i, x_j] \Vdash \varphi_{inc}^k$. $\qquad\square$

Formulae φ_{eq}^i and φ_{inc}^i are used to place the y coordinate of the points of the corridor on the correct u-interval. Let φ_y be defined as follows:

$$\varphi_y = \begin{cases} \langle A \rangle \big(u \wedge \bigwedge_{0 \le i < n} \neg y_i \big) \wedge \\ [G] \bigwedge_{0 \le i < n} \Big((y_i \leftrightarrow [B]y_i) \wedge (\neg y_i \leftrightarrow [B]\neg y_i) \Big) \wedge \\ [G] \big(u \to \varphi_{inc}^0 \big) \end{cases}$$

Lemma 7. *Let $M, [x, y] \Vdash \varphi_{u-chain} \wedge \varphi_y$ and let $x_0, x_1, \ldots$ be the sequence of points of Lemma 5. Then, for every interval $[x_i, z]$, it holds that $\text{bin}([x_i, z], 0) = i \mod 2^n$.*

Proof. We prove the thesis by induction on i. As for the base case, consider an interval $[x_0, z]$, with $x_0 = y$. By the first conjunct of φ_y, it holds that $[x_0, x_1] \Vdash \neg y_j$, for each $j \in \{0, \ldots n - 1\}$. By the second conjunct, for all z, it holds that $[x_0, z] \Vdash \neg y_j$, for each j, and thus $\text{bin}([x_0, z], 0) = 0$. As for the inductive step, let $i > 0$ and let the thesis hold for the interval $[x_{i-1}, x_i]$. By the third conjunct of φ_y, it holds that $[x_{i-1}, x_i] \Vdash \varphi_{inc}^0$. Then, by Lemma 6, $\text{bin}([x_i, x_{i+1}], 0) = \text{bin}([x_{i-1}, x_i], 0) + 1 \mod 2^n$, and thus, by the inductive hypothesis, $\text{bin}([x_i, x_{i+1}], 0) = i \mod 2^n$. By the second conjunct of φ_y, we can conclude that $\text{bin}([x_i, z], 0) = \text{bin}([x_i, x_{i+1}], 0)$ for all $z > x_i$. $\qquad\square$

Now, we can use u-intervals to represent the tiling function $f : \mathbb{N} \times \{0, \ldots, 2^n - 1\} \mapsto T$. Let $T = \{t_1, \ldots, t_k\}$ be the set of tiles. We first associate a unique tile $f(p)$ with each point p in the corridor as follows:

$$\varphi_f = [G] \big(u \leftrightarrow \bigvee_{1 \le i \le k} t_i \big) \wedge [G] \bigwedge_{1 \le i < j \le k} \neg(t_i \wedge t_j).$$

Then, we enforce the constraints (i)–(iv) on the tiling function:

$$\varphi_\perp = [G]\Big(u \wedge \bigwedge_{0 \le i < n} \neg y_i \to t_\perp\Big);$$

$$\varphi_\top = [G]\Big(u \wedge \bigwedge_{0 \le i < n} y_i \to t_\top\Big);$$

$$\varphi_H = [G] \bigwedge_{1 \le i \le k} \Big((\varphi_{eq}^0 \wedge [B](\langle A \rangle u \to \neg \varphi_{eq}^0) \wedge \langle B \rangle t_i) \to \bigvee_{(t_i, t_j) \in H} \langle A \rangle t_j\Big);$$

$$\varphi_V = [G] \bigwedge_{1 \le i \le k} \Big(t_i \to \bigvee_{(t_i, t_j) \in V} \langle A \rangle t_j\Big).$$

The formula encoding the tiling problem is:

$$\varphi_{\mathcal{T}} = \varphi_{u-chain} \wedge \varphi_y \wedge \varphi_f \wedge \varphi_\perp \wedge \varphi_\top \wedge \varphi_H \wedge \varphi_V.$$

Lemma 8. *Let $M, [x, y] \Vdash \varphi_{\mathcal{T}}$ and let $x_0, x_1, \ldots$ be the sequence of points of Lemma 5. Then, for every interval $[x_i, x_{i+1}]$, it holds that:*

    (i) there exists a unique tile t_j such that $[x_i, x_{i+1}] \Vdash t_j$;
    (ii) if $i \mod 2^n = 0$, then $[x_i, x_{i+1}] \Vdash t_\perp$;
    (iii) if $i \mod 2^n = 2^n - 1$, then $[x_i, x_{i+1}] \Vdash t_\top$;
    (iv) if $[x_i, x_{i+1}] \Vdash t_j$ and $[x_{i+2^n}, x_{i+2^n+1}] \Vdash t_m$, then $(t_j, t_m) \in H$;
    (v) if $[x_i, x_{i+1}] \Vdash t_j$ and $[x_{i+1}, x_{i+2}] \Vdash t_m$, then $(t_j, t_m) \in V$.

Proof. Let $M, [x, y] \Vdash \varphi_{\mathcal{T}}$ and let $x_0 < x_1 < x_2 < \ldots$ be the infinite ascending sequence whose existence has been proved by Lemma 5. Item (i) immediately follows from φ_f. To prove (ii), let $[x_i, x_{i+1}]$ be a u-interval such that $i \mod 2^n = 0$. By Lemma 7, $\text{bin}([x_i, x_{i+1}], 0) = 0$ and thus $[x_i, x_{i+1}] \Vdash \bigwedge_{0 \le j < n} \neg y_j$. By $\varphi_\perp$, we can conclude that $[x_i, x_{i+1}] \Vdash t_\perp$. Item (iii) can be proved as (ii). To prove (iv), suppose, by contradiction, that there exist two intervals $[x_i, x_{i+1}]$ and $[x_{i+2^n}, x_{i+2^n+1}]$ such that $[x_i, x_{i+1}] \Vdash t_j$ and $[x_{i+2^n}, x_{i+2^n+1}] \Vdash t_m$, but $(t_j, t_m) \notin H$. Consider now the interval $[x_i, x_{i+2^n}]$. Since $i \mod 2^n = i + 2^n \mod 2^n$, by Lemma 7, it holds that $\text{bin}([x_i, x_{i+2^n}], 0) = \text{bin}([x_{i+2^n}, x_{i+2^n+1}], 0)$. By Lemma 6, we have that $[x_i, x_{i+2^n}] \Vdash \varphi_{eq}^0$. Moreover, since for each $j \in \{i + 1, \ldots, i + 2n - 1\}$, it holds that $\text{bin}([x_j, x_{j+1}], 0) = j \mod 2^n \ne i \mod 2^n = \text{bin}([x_i, x_j], 0)$, by Lemma 6, we have that $[x_i, x_j] \Vdash \neg \varphi_{eq}^0$, for each $j \in \{i + 1, \ldots, i + 2n - 1\}$, and thus $[x_i, x_j] \Vdash [B](\langle A \rangle u \to \neg \varphi_{eq}^0)$. Now, since $[x_i, x_{i+1}] \Vdash t_j$, by φ_H, it follows that $(t_j, t_m) \in H$, against the hypothesis that $(t_j, t_m) \notin H$. Finally, (v) immediately follows from φ_V. $\qquad\square$

The next lemma proves the correctness of the encoding.

Lemma 9. *The* AB*-formula $\varphi_{\mathcal{T}}$ is satisfiable if and only if $\mathcal{T} = (T, t_{\perp}, t_{\top}, H, V, n)$ is a satisfiable instance of the exponential-corridor tiling problem.*

Proof. Let us first assume $\varphi_{\mathcal{T}}$ to be satisfiable, that is, there are an interval model M and an interval $[x, y]$ such that $M, [x, y] \Vdash \varphi_{\mathcal{T}}$. By Lemmas 5–8, we can easily check that M represents a satisfiable instance of $\mathcal{T}$.

To prove the converse direction, let $\mathcal{T} = (T, t_{\perp}, t_{\top}, H, V, n)$ be a satisfiable instance of the tiling problem and let $f : \mathbb{N} \times \{0, \ldots, 2^n - 1\} \mapsto T$ be the corresponding tiling function. A model over $\mathbb{Q}$ that encodes f can be obtained as follows. First, we build the infinite chain of u-intervals $[0, 1], [1, 2], [2, 3], \ldots$. Then, we define the enumeration $g(x, y) = x \cdot 2^n + y$ that associates a unique natural number with every point $p = (x, y) \in \mathbb{N} \times \{0, \ldots, 2^n - 1\}$. For every point (x, y), we put the interval $[g(x, y), g(x, y) + 1]$ in the valuation of the proposition letter corresponding to the tile $f(x, y)$. It is easy to define a valuation for all the other proposition letters that conforms to their intended meaning, as shown in the above encoding, thus obtaining a model M such that $M, [-1, 0] \Vdash \varphi_{\mathcal{T}}$. $\qquad\square$

A similar reduction can be given for the fragment $A\overline{B}$.

Theorem 4. *The satisfiability problem for* AB *and* $A\overline{B}$ *over* $\mathbb{Q}$ *is* EXPSPACE-*hard.*

7. Non-primitive recursive fragments

To complete the picture given in Figure 2, we just need to add one missing piece concerning the non-primitive recursive fragments. The non-primitive recursiveness of $A\overline{A}B$ and $A\overline{A}\overline{B}$ has been shown in [15]. Here, we prove that, in fact, all fragments containing $\overline{A}B$ or $\overline{A}\overline{B}$ are non-primitive recursive. In the following, we shall make an essential use of lossy counter machines. They can be viewed as a variant of Minsky counter automata where transitions may non-deterministically decrease the values of counters. A comprehensive survey on lossy/faulty machines and the relevant complexity, decidability, and undecidability results can be found in [29].

Formally, a (Minsky) *counter automaton* is a tuple $\mathcal{A} = (Q, q_0, C, \Delta)$, where Q is a finite set of control states, $q_0 \in Q$ is the initial state, $C = \{c_1, \ldots, c_k\}$ is a finite set of counters, whose values range over $\mathbb{N}$, and $\Delta \subseteq$

$Q \times L \times Q$, with $L = \{inc, dec, ifz\} \times \{1, \ldots, k\}$ (instruction set), is the transition relation. A *configuration* of $\mathcal{A}$ is a pair $(q, \bar{v})$, where $q \in Q$ and $\bar{v} \in \mathbb{N}^k$ is a vector of counter values (we denote the i-th component of $\bar{v}$ by v_i). There is a *transition* from configuration $(q, \bar{v})$ to configuration $(q', \bar{v}')$ via instruction l, $(q, \bar{v}) \xrightarrow{l} (q', \bar{v}')$ for short, if, and only if, $(q, l, q') \in \Delta$ and either *(i)* $l = (inc, i)$, $v_i' = v_i + 1$, and $v_j' = v_j$ for each $j \neq i$ (*increment*), or *(ii)* $l = (dec, i)$, $v_i' = v_i - 1$, and $v_j' = v_j$ for each $j \neq i$ (*decrement*), or *(iii)* $l = (ifz, i)$, $v_i = 0$, and $v_j' = v_j$ for all j (*test*). A *run* is a finite or infinite sequence of configurations such that, for every pair of consecutive configurations $(q, \bar{v}), (q', \bar{v}')$, $(q, \bar{v}) \xrightarrow{l} (q', \bar{v}')$ for some l. *Lossy counter machines* differ from Minsky counter automata for the possible presence of faulty transitions in a run, i.e., transitions that nondeterministically decrease the value of some counter. Formally, we have a *faulty transition* $(q, \bar{v}) \xrightarrow{l}_{\dagger} (q', \bar{v}')$ if, and only if, there is a transition $(q, \bar{v}_{\dagger}) \xrightarrow{l} (q', \bar{v}_{\dagger}')$ with both $\bar{v} \geq \bar{v}_{\dagger}$ and $\bar{v}_{\dagger}' \geq \bar{v}'$, where the ordering $\geq$ is defined component-wise in the obvious way. The *non-termination problem* for lossy machines is the problem of deciding whether a lossy machine $\mathcal{A}$ has at least one infinite run starting from the *initial configuration* $(q_0, \bar{0})$. In [29], it has been shown that this problem is non-primitive recursive. In the following, we reduce the non-termination problem for lossy counter machines to the satisfiability problem for $\overline{\mathsf{A}}\mathsf{B}$ over $\mathbb{Q}$.

Let $\mathcal{A} = (Q, q_0, C, \Delta)$ be a lossy counter machine. We build an $\overline{\mathsf{A}}\mathsf{B}$-formula $\varphi_{\mathcal{A}}$ which turns out to be satisfiable over $\mathbb{Q}$ if and only if $\mathcal{A}$ has at least one infinite run starting from the initial configuration. The infinite run is encoded (left-to-right) to the left of the interval $[x, y]$ where the formula is evaluated by exploiting the fact that, for all $x' < x$, there are infinitely many intervals between x' and x. We shall make use of the proposition letters u for *units*, $q_0, \ldots, q_{|Q|-1}$ for *states*, *conf* for *configurations*, $c_1 \ldots, c_{|C|}$ for *counters*, and *corr*, *corr$_1$*, $\ldots$, *corr$_{|C|}$* for *correspondences*. Proposition letters $c_1 \ldots, c_{|C|}$ allow us to encode the counters of $\mathcal{A}$: given a configuration where the value of the i-th counter is n, the corresponding *conf*-interval will contain precisely n c_i-intervals (for every proposition letter p, we call p-intervals those intervals that satisfy p). Additional proposition letters will be exploited in the reduction for technical reasons.

Given an interval $[x, y]$, let $\mathcal{G}_{[x,y]}^{\overline{\mathsf{A}}\mathsf{B}}$ be the set of intervals that are reachable from $[x, y]$ by $\overline{\mathsf{A}}\mathsf{B}$-formulae, that is, $\mathcal{G}_{[x,y]}^{\overline{\mathsf{A}}\mathsf{B}} = \{[x, y]\} \cup \{[x, z] \mid z < y\} \cup \{[w, z] \mid z \leq x\}$. In the rest of the section, even when not explicitly stated, we focus

on intervals in $\mathcal{G}^{\overline{\mathsf{A}}\mathsf{B}}_{[x,y]}$. Let $[G^{\overline{\mathsf{A}}\mathsf{B}}]\varphi$ be the following shorthand: $[G^{\overline{\mathsf{A}}\mathsf{B}}]\varphi = \varphi \wedge [B]\varphi \wedge [\overline{A}]\varphi \wedge [A][\overline{A}]\varphi$. Modality $[G^{\overline{\mathsf{A}}\mathsf{B}}]$ can be thought of as the global modality for the fragment $\overline{\mathsf{A}}\mathsf{B}$, as it allows one to predicate about all intervals in the relevant set $\mathcal{G}^{\overline{\mathsf{A}}\mathsf{B}}_{[x,y]}$. This is formally stated by the following proposition, whose proof is straightforward and thus omitted (as we did in the previous section, to keep the notation light, and with an abuse of notation, from now on we omit the superscript $\overline{\mathsf{A}}\mathsf{B}$ in $[G^{\overline{\mathsf{A}}\mathsf{B}}]$ and denote it simply by $[G]$).

Proposition 2. *Let φ be an $\overline{\mathsf{A}}\mathsf{B}$-formula and $[x, y]$ be an interval. Then, $M, [x, y] \Vdash [G]\varphi$ if and only if $M, [w, z] \Vdash \varphi$, for each $[w, z] \in \mathcal{G}^{\overline{\mathsf{A}}\mathsf{B}}_{[x,y]}$.*

As a preliminary step, for a given proposition letter p, we show how to prevent p-intervals from being proper prefixes of or to start inside other p-intervals. To this end, we use an auxiliary proposition letter p_b (the subscript 'b' stands for 'beginning'). Let *no_overlap(p)* be a shorthand for the formula:

$$[G](p \rightarrow \neg\langle B\rangle p) \wedge [G](p \rightarrow [B]p_b) \wedge [G](p \rightarrow \neg\langle \overline{A}\rangle p_b).$$

It can easily shown that the following proposition holds.

Proposition 3. *If $M, [x, y] \Vdash$ no_overlap(p), then no pair of distinct p-intervals $[w, z], [w', z'] \in \mathcal{G}^{\overline{\mathsf{A}}\mathsf{B}}_{[x,y]}$ is such that $w \leq w' < z$.*

The first step in the construction of the formula $\varphi_{\mathcal{A}}$ is the *discretization* of the domain, which makes an essential use of the proposition letter u. Let $\varphi_{u-chain}$ be the formula:

$$\varphi_{u-chain} = \begin{cases} \langle\overline{A}\rangle\langle\overline{A}\rangle(u \wedge start) \wedge [\overline{A}](\langle\overline{A}\rangle u \rightarrow \langle B\rangle u) \wedge no\_overlap(u) \wedge \\ [G](start \rightarrow [\overline{A}]\neg u \wedge [\overline{A}][\overline{A}]\neg u) \wedge [G](start \rightarrow u) \end{cases}$$

The next lemma shows that, when evaluated over an interval $[x, y]$, $\varphi_{u-chain}$ generates an infinite chain of u-intervals, which starts with a *start*-interval and is confined to the left of $[x, y]$.

Lemma 10. *Let $M, [x, y] \Vdash \varphi_{u-chain}$. Then, M features an infinite sequence of points $x_0, x_1, \ldots$, with $x_0 < x_1 < \ldots < x_i < \ldots < x < y$, such that:*

(i) *for every $i \geq 0$, $[x_i, x_{i+1}]$ is a u-interval;*

(ii) *no other u-interval $[w, z]$ exists with $z \leq x_i$ for some $i \geq 0$;*

(iii) *$[x_0, x_1]$ is a start-interval and no other start-interval $[w, z]$ exists with $z \leq x_i$ for some $i \geq 0$.*

Proof. The first conjunct of $\varphi_{u-chain}$ guarantees the existence of a u-interval, say it $[x_0, x_1]$, with $x_1 < x$, which is a *start*-interval as well. The second conjunct states that every u-interval $[w, z]$, with $z < x$, is immediately followed by another u-interval $[z, z']$, with $z' < x$. Thus, (i) holds. By way of contradiction, assume that there exists a u-interval $[w, z]$, with $z \leq x_i$ for some $i \geq 0$, such that $[w, z] \neq [x_j, x_{j+1}]$ for every $j \geq 0$. We must distinguish two cases. First, if $z \leq x_0$, then there would be a *start*-interval preceded by a u-interval, thus contradicting the fourth conjunct of $\varphi_{u-chain}$. Otherwise, if $x_i \leq z < x_{i+1}$ for some $i \geq 0$, then, by the second conjunct of $\varphi_{u-chain}$, there would be a u-interval (different from $[x_i, x_{i+1}]$) that starts at z and ends at some point to the left of x, which contradicts *no_overlap*(u). Therefore, (ii) is proved. Finally, we already proved the existence of the *start*-interval $[x_0, x_1]$, which is also a u-interval, and, now, we show that there are no other *start*-intervals $[w, z]$, with $z \leq x_i$ for some $i \geq 0$. By way of contradiction, assume that there exists one such interval $[w, z]$. By the last conjunct of $\varphi_{u-chain}$, $[w, z]$ is also a u-interval. By item (ii) above, $[w, z] = [x_i, x_{i+1}]$ for some i, which leads to contradiction with the fourth conjunct of $\varphi_{u-chain}$. $\square$

The sequence of u-intervals is used to encode a run of the machine: each u-interval corresponds to either a state (q_i-intervals) or a counters' instance (c_i-intervals); states and counters' instances will be then suitably wrapped into configurations. To handle increment and decrement transitions, we introduce the proposition letters c_i^+ or c_i^- ($1 \leq i \leq |C|$), respectively, that correspond to special counters' instances. Let $\mathcal{QC} = \{q_i \mid 0 \leq i \leq |Q| - 1\} \cup \{c_i \mid 1 \leq i \leq |C|\}$ be the set of proposition letters for states and counters' instances. The correct behaviour of these proposition letters, as well as of proposition letters c_i^+ and c_i^-, is forced by the following formula:

$$\varphi_{units} = \begin{cases} [G](u \leftrightarrow \bigvee_{p \in \mathcal{QC}} p) \wedge [G] \bigwedge_{p_1, p_2 \in \mathcal{QC}, p_1 \neq p_2}(p_1 \rightarrow \neg p_2) \wedge \\ [G] \bigwedge_{i \in \{1, \ldots, |C|\}} \left((c_i^- \rightarrow c_i) \wedge (c_i^+ \rightarrow c_i)\right) \end{cases}$$

The next formula forces the existence of an infinite sequence of configurations (*conf*-intervals). The first configuration is associated with the u-interval $[x_0, x_1]$ (*start*-interval), and it contains the initial state q_0 only. This conforms to the requirement that the value of each counter in the initial configuration is 0. Moreover, we guarantee that (i) the endpoints of configurations are also endpoints of u-intervals and (ii) every configuration contains a state, which

45

is associated with its last u-interval.

$$\varphi_{conf-chain} = \begin{cases} [G](start \to conf \wedge q_0) \wedge [\overline{A}](\langle \overline{A} \rangle conf \to \langle B \rangle conf) \wedge \\ no\_overlap(conf) \wedge [G](start \to [\overline{A}]\neg conf \wedge [A][\overline{A}]\neg conf) \wedge \\ [G](conf \to (u \vee \langle B \rangle u)) \wedge [G](\langle \overline{A} \rangle conf \leftrightarrow \langle \overline{A} \rangle \bigvee_{i=0,..,|Q|-1} q_i) \end{cases}$$

Lemma 11. *Let $M, [x, y] \Vdash \varphi_{u-chain} \wedge \varphi_{units} \wedge \varphi_{conf-chain}$ and let $x_0, x_1, \ldots$ be the sequence of points of Lemma 10. Then, M features an infinite sequence of points $x'_0, x'_1, \ldots$, with $x'_0 = x_0 < x'_1 = x_1 < x'_2 < x'_3 < \ldots < x'_i < \ldots < x < y$, such that:*

 (i) *$[x'_0, x'_1]$ is a q_0-interval and, for all $i \geq 0$, $[x'_i, x'_{i+1}]$ is a conf-interval;*
 (ii) *no other conf-interval $[w, z]$ exists with $z \leq x'_i$ for some $i \geq 0$;*
 (iii) *for all $i \geq 0$, there exist x', x'', and j, with $x' \leq x'_{i+1}$, $x'' \geq x'_i$, and $0 \leq j \leq |Q| - 1$, such that $[x'_i, x']$ is a u-interval and $[x'', x'_{i+1}]$ is a q_j-interval;*
 (iv) *for all $i \geq 0$, no q_j-interval $[w, z]$ exists with $x'_i < z < x'_{i+1}$.*

Proof. By the first conjunct of $\varphi_{conf-chain}$, $[x'_0, x'_1]$ $(= [x_0, x_1])$ is a q_0-interval. Thanks to the second conjunct of $\varphi_{conf-chain}$, the existence of an infinite sequence of *conf*-intervals can be proved by exploiting the same argument used in the proof of item (i) of Lemma 10. This proves (i). To prove (ii), we can proceed as for item (ii) of Lemma 10, by using the third and fourth conjunct of $\varphi_{conf-chain}$. Let us consider now (iii). By the fifth conjunct of $\varphi_{conf-chain}$, every *conf*-interval $[x'_i, x'_{i+1}]$ is either a u-interval or it is started by a u-interval, and in either case it follows that there exists $x' \leq x'_{i+1}$ such that $[x'_i, x']$ is a u-interval. Moreover, by the sixth conjunct of $\varphi_{conf-chain}$, there exist x'' and j, with $x'' < x'_{i+1}$ and $0 \leq j \leq |Q| - 1$, such that $[x'', x'_{i+1}]$ is a q_j-interval. By the first conjunct of φ_{units}, $[x'', x'_{i+1}]$ is a u-interval as well. To complete the proof, we need to show that $x'' \geq x'_i$. By way of contradiction, let us assume that $x'' < x'_i$. By the fifth conjunct of $\varphi_{conf-chain}$, there is a u-interval starting at x'_i, and, since $x'' < x'_i < x'_{i+1}$, it starts inside the u-interval $[x'', x'_{i+1}]$, thus violating the third conjunct of $\varphi_{u-chain}$. So, (iii) holds. Finally, as for (iv), assume, by way of contradiction, the existence of two indexes i and j such that $[w, z]$ is a q_j-interval and $x'_i < z < x'_{i+1}$. By the sixth conjunct of $\varphi_{conf-chain}$, there is a *conf*-interval ending at z, which contradicts item (ii) of the lemma. $\square$

The formula built so far constrains every *conf*-interval to contain exactly one q_j-interval (one state), all the other u-intervals being counters' instances.
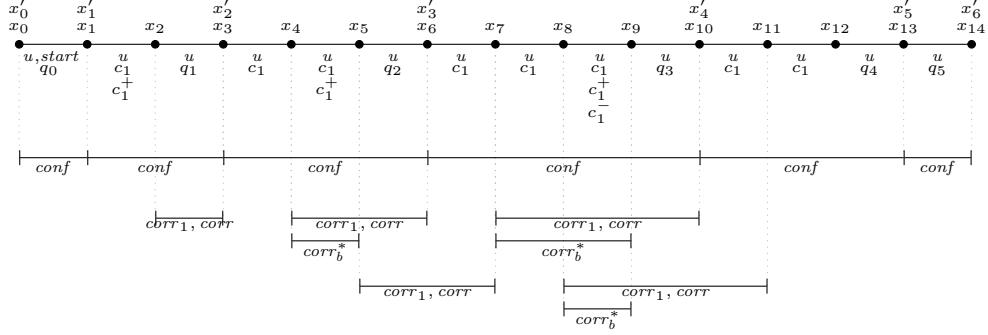
Figure 5: A (fragment of a) model satisfying $\varphi_{\mathcal{A}}$ (for the sake of readability, we omit the symbol $corr_b$ that labels all the prefixes of $corr$ intervals).

We now show how to force, for each $1 \leq i \leq |C|$, the number of counters' instances (c_i-intervals) in every $conf$-interval to be equal to the value of the i-th counter in the corresponding configuration of the run being encoded.

Here the differences between a Minsky counter automaton and a lossy counter machine come into play. In the former, the values of the counters are univocally determined by the transition: the values of the counters affected by the transition operation must change accordingly, while the ones of the others remain unchanged. In the latter, it suffices to guarantee that no counter is ever incremented, with the possible exception of the ones involved in an increment operation, and to force the value of the counters involved in a decrement operation to be decreased at least by one. It is worth noticing that a counter's value can possibly decrease even in case it is involved in an increment operation, due to a faulty behaviour of the machine.

To encode these constraints, we make use of the additional proposition letters $corr, corr_1, \ldots, corr_{|C|}, corr_b$, and $corr_b{}^*$, besides proposition letters u, q_i, with $i \in \{0, \ldots, |Q| - 1\}$, and c_j, c_j^+, c_j^-, with $j \in \{1, \ldots, |C|\}$. The idea is that c_i^+-intervals (resp., c_i^--intervals) are used to encode increment (resp., decrement) transitions, while $corr_i$-intervals maintain correspondence between c_i-intervals in adjacent configurations, in order to prevent the addition of c_i-intervals unless an increment transition occurs (see Figure 5).

More precisely, every c_i-interval $[x, y]$ that is not labeled with c_i^+ corresponds to a c_i-interval $[w, z]$ in the previous configuration. Such a correspondence is encoded by labeling the interval $[z, x]$ with $corr_i$. On the contrary, c_i-intervals that are also labeled with c_i^+ do not correspond to any c_i-interval

(in the previous configuration) because they are added to encode an incre-ment transition. Similarly, a c_i-interval that is also labeled with c_i^- does not correspond to any c_i-interval in the next configuration: it disappears in order to simulate a decrement of the i-th counter. Notice that a c_i^+-interval occurring in a configuration encodes an increment transition *leading* to that configuration, while a c_i^--intervals occurring in a configuration encodes a decrement transition *starting* from that configuration. Finally, proposition letters $corr_b$ and $corr_b{}^*$ are used to force $corr$-intervals to connect intervals belonging to consecutive configurations.

The above conditions are encoded by the following formula.

$$\varphi_{corr} = \begin{cases} [G] \bigwedge_{i=1,\ldots,|C|}((c_i \wedge \neg c_i^+) \rightarrow \langle \overline{A} \rangle corr_i) \wedge \\ [G] \bigwedge_{i=1,\ldots,|C|}(corr_i \rightarrow \langle \overline{A} \rangle(c_i \wedge \neg c_i^-)) \wedge \\ [G] \bigwedge_{i=1,\ldots,|C|}(corr_i \rightarrow corr) \wedge [G](\langle \overline{A} \rangle corr \rightarrow \langle \overline{A} \rangle u) \wedge \\ [G](corr \rightarrow (\bigvee_{i=0,\ldots,|Q|-1}(q_i \vee \langle B \rangle q_i) \vee \langle B \rangle corr_b{}^*)) \wedge \\ [G](\langle \overline{A} \rangle corr_b{}^* \rightarrow \langle \overline{A} \rangle u) \wedge \\ [G]((u \wedge \bigwedge_{i=0,\ldots,|Q|-1} \neg q_i) \rightarrow \neg \langle \overline{A} \rangle corr_b{}^*) \wedge \\ [G](corr \rightarrow [B] corr_b) \wedge \\ [G]((\bigvee_{i=0,\ldots,|Q|-1} q_i) \rightarrow [\overline{A}](corr_b \rightarrow corr_b{}^*)) \wedge \\ [G] \bigwedge_{i=0,\ldots,|Q|-1}(corr_b{}^* \rightarrow \neg q_i \wedge [B] \neg q_i) \wedge \\ [G](corr_b{}^* \rightarrow [B] \neg corr_b{}^*) \wedge [G](corr \rightarrow [B] \neg corr) \end{cases}$$

The correctness of the encoding is formally stated by the next two lemmas.

Lemma 12. *Let $M, [x, y] \Vdash \varphi_{u-chain} \wedge \varphi_{units} \wedge \varphi_{conf-chain} \wedge \varphi_{corr}$ and let $x'_0, x'_1, \ldots$ be the sequence of points of Lemma 11. Then, for each $1 \leq i \leq |C|$, it holds that:*

(i) *for every c_i-interval $[w, z]$, which is not a c_i^+-interval, there exists $w' < w$ such that $[w', w]$ is a $corr_i$-interval;*

(ii) *for every $corr_i$-interval $[w, z]$, there exists $w' < w$ such that $[w', w]$ is a c_i-interval, but not a c_i^--interval;*

(iii) *for every $corr_i$-interval $[w, z]$, if $x'_{j+1} \leq z < x'_{j+2}$, for some $j \geq 0$, then $x'_j < w < x'_{j+1}$, that is, the conf-interval the right endpoint of a $corr_i$-interval belongs to is the successor of the conf-interval its left endpoint belongs to.*

Proof. Item (i) directly follows from the first conjunct of φ_{corr}, and item (ii) is an immediate consequence of the second one.

To prove item (iii), by way of contradiction, let $[w, z]$ be a $corr_i$-interval such that $x'_{j+1} \leq z < x'_{j+2}$, for some $j \geq 0$, and either $w \geq x'_{j+1}$ or $w \leq x'_j$. Notice that, by the third conjunct of φ_{corr}, $[w, z]$ is a $corr$-interval as well.

Let us assume that $w \geq x'_{j+1}$. By item (iii) of Lemma 11, there is $\bar{x}_{j+2}$ such that $[\bar{x}_{j+2}, x'_{j+2}]$ is a q_k-interval, for some $0 \leq k \leq |Q| - 1$. By the third conjunct of φ_{corr}, there is a u-interval ending in z and, by Lemma 10, there is also a u-interval starting at z. Therefore, it holds $z \leq \bar{x}_{j+2}$ (otherwise, there would be two overlapping u-intervals). Since $x'_{j+1} \leq w < z \leq \bar{x}_{j+2} < x'_{j+2}$, by item (iv) of Lemma 11, neither $[w, z]$ nor its prefixes are q_m-intervals, for any $0 \leq m \leq |Q| - 1$. Thus, the fourth conjunct of φ_{corr} implies the existence of a $corr_b^*$-interval $[w, z']$, with $w < z' < z$. By the fifth conjunct of φ_{corr}, there is a u-interval ending in z' and, by Lemma 10, there is also a u-interval starting at z' and ending not later z (otherwise, there would be two overlapping u-intervals). Let $[z', z'']$ be such an interval. Once again, since $x'_{j+1} \leq w < z' < z'' \leq z < x'_{j+2}$, by item (iv) of Lemma 11, $[z', z'']$ is not a q_m-interval, for any $0 \leq m \leq |Q| - 1$. Hence, $[z', z''] \Vdash (u \wedge \bigwedge_{i=0,\ldots,|Q|-1} \neg q_i) \wedge \langle \overline{A} \rangle corr_b^*$, thus contradicting the sixth conjunct of φ_{corr}.

Now, suppose that $w \leq x'_j$. Two cases are possible. If $j = 0$, then, by item (iii) of Lemma 10 and by Lemma 11, $[x'_j, x'_{j+1}] = [x_j, x_{j+1}]$ is the only $start$-interval, and there is no u-interval to the left of it, thus contradicting item (ii) above. Let us assume $j > 0$. By item (iii) of Lemma 11, there are two points $\bar{x}_j \geq x'_{j-1}$ and $\bar{x}_{j+1} \geq x'_j$ such that $[\bar{x}_j, x'_j]$ and $[\bar{x}_{j+1}, x'_{j+1}]$ are, respectively, a $q_{k'}$-interval and a $q_{k''}$-interval, for some $0 \leq k', k'' \leq |Q| - 1$. By item (ii) above, $w \leq \bar{x}_j$ (otherwise, there would be two overlapping u-intervals). By the seventh conjunct of φ_{corr}, $[w, \bar{x}_{j+1}]$ is a $corr_b$-interval and, by the eighth conjunct of φ_{corr}, it is a $corr_b^*$-interval as well. We distinguish three cases. If $w = \bar{x}_j$ and $x'_j = \bar{x}_{j+1}$, then $[\bar{x}_j, x'_j]$ is both a $q_{k'}$- and a $corr_b^*$-interval, thus violating the ninth conjunct of φ_{corr}. If $w = \bar{x}_j$ and $x'_j < \bar{x}_{j+1}$, then the $q_{k'}$-interval $[\bar{x}_j, x'_j]$ starts the $corr_b^*$-interval $[w, \bar{x}_{j+1}]$, thus violating again the ninth conjunct of φ_{corr}. If $w < \bar{x}_j$, then, by the seventh conjunct of φ_{corr}, $[w, \bar{x}_j]$ is a $corr_b$-interval and, by the eighth conjunct of φ_{corr}, it is a $corr_b^*$-interval as well, and thus the $corr_b^*$-interval $[w, \bar{x}_j]$ starts the $corr_b^*$-interval $[w, \bar{x}_{j+1}]$, thus violating the tenth conjunct of φ_{corr}. $\qquad \square$

Lemma 13. *Let $M, [x, y] \Vdash \varphi_{u-chain} \wedge \varphi_{units} \wedge \varphi_{conf-chain} \wedge \varphi_{corr}$. For every pair of consecutive conf-intervals $[x'_j, x'_{j+1}]$ and $[x'_{j+1}, x'_{j+2}]$, with $j \geq 0$, where, for $1 \leq i \leq |C|$, the number of c_i-intervals in $[x'_j, x'_{j+1}]$ and $[x'_{j+1}, x'_{j+2}]$ is respectively n^i_j and n^i_{j+1}, it holds that:*

(i) *if* $[x'_{j+1}, x'_{j+2}]$ *does not contain* c_i^+*-intervals, then* $n^i_{j+1} \leq n^i_j$;

(ii) *if* $[x'_{j+1}, x'_{j+2}]$ *contains (exactly) one* c_i^+*-interval, then* $n^i_{j+1} \leq n^i_j + 1$;

(iii) *if* $[x'_j, x'_{j+1}]$ *contains a* c_i^-*-interval and* $[x'_{j+1}, x'_{j+2}]$ *does not contain* c_i^+*-intervals, then* $n^i_{j+1} \leq n^i_j - 1$.

Proof. For each c_i-interval $[w, z]$ contained in $[x'_{j+1}, x'_{j+2}]$ that is not a c_i^+-interval, *(a)* there exists a $corr_i$-interval $[w', w]$ (Lemma 12, item (i)), *(b)* there exists a c_i-interval ending in w' (Lemma 12, item (ii)), and *(c)* such a c_i-interval is contained in $[x'_j, x'_{j+1}]$ (Lemma 12, item (iii)). Thus, for each c_i-interval in $[x'_{j+1}, x'_{j+2}]$ there is a corresponding c_i-interval in $[x'_j, x'_{j+1}]$ (unless such a c_i-interval is also a c_i^+-interval). Moreover, by the tenth conjunct of φ_{corr} (in particular, $[G](corr \rightarrow [B]\neg corr)$), two distinct c_i-intervals in $[x'_{j+1}, x'_{j+2}]$ cannot be linked to the same c_i-interval in $[x'_j, x'_{j+1}]$. Hence, if $[x'_{j+1}, x'_{j+2}]$ does not contain c_i^+-intervals, then $n^i_{j+1} \leq n^i_j$, and (i) holds. As for (ii), the proof is basically the same. We only need to observe that there might be a c_i-interval in $[x'_{j+1}, x'_{j+2}]$, namely, the one which is also a c_i^+-interval, that has no corresponding c_i-interval in $[x'_j, x'_{j+1}]$ (due to the first conjunct of φ_{corr}). The thesis immediately follows. A similar argument works for (iii). We only need to observe that there must be at least one c_i-interval in $[x'_j, x'_{j+1}]$, namely, the one which is also a c_i^--interval, that does not correspond to any c_i-interval in $[x'_{j+1}, x'_{j+2}]$ (due to the second conjunct of φ_{corr}). The thesis immediately follows. $\square$

To encode the transitions of the run, we introduce an auxiliary proposition letter $conf_{(q,l,q')}$ for each $(q, l, q') \in \Delta$. These proposition letters hold on *conf*-intervals and carry information on which transition produced the corresponding configuration. To start with, every configuration, but *start*, is the result of exactly one transition, as formally stated by the formula:

$$\varphi_{conf} = \begin{cases} [G]((conf \wedge \neg start) \leftrightarrow (\bigvee_{(q,l,q')\in\Delta} conf_{(q,l,q')})) \wedge \\ [G] \bigwedge_{(q,l,q')\in\Delta}(conf_{(q,l,q')} \rightarrow (\bigwedge_{(q'',l',q''')\neq(q,l,q')} \neg conf_{(q'',l',q''')})) \end{cases}$$

The next four formulae encode the effects of transitions by imposing suitable constraints on *conf*-intervals on the basis of the transitions that generate them. The first formula φ_{inc} forces every configuration produced by an increment transition (i being the involved counter) to contain at most one c_i^+-interval (the first three conjuncts of φ_{inc}). In order to do that, the new auxiliary proposition letter $c_{i,b}^+$ is used to label intervals whose right endpoint coincides with the right endpoint of a c_i^+-interval (first conjunct of φ_{inc}), and

whose left endpoint coincides with the left endpoint of the *conf*-interval to which the relevant c_i^+-interval belongs (second conjunct of φ_{inc}); then, by imposing that a $c_{i,b}^+$-interval cannot be a prefix of another one (third conjunct of φ_{inc}), we are guaranteed that a configuration contains at most one c_i^+-interval. Notice that a configuration produced by an increment transition may possibly contain no c_i^+-intervals at all, due to a faulty behaviour of the machine that suppressed the effect of the increment. Moreover, c_i^+-intervals are allowed only within configurations produced by increment transitions (last conjunct).

$$\varphi_{inc} = \begin{cases} [G] \bigwedge_{i=1,\dots,|C|}(\langle \overline{A} \rangle c_{i,b}^+ \leftrightarrow \langle \overline{A} \rangle c_i^+) \wedge \\ [G] \bigwedge_{i=1,\dots,|C|}(c_{i,b}^+ \to (\langle \overline{A} \rangle conf \wedge [B] \neg conf)) \wedge \\ [G] \bigwedge_{i=1,\dots,|C|}(c_{i,b}^+ \to [B] \neg c_{i,b}^+) \wedge \\ [G] \bigwedge_{i=1,\dots,|C|}((conf \wedge \langle B \rangle c_{i,b}^+) \to \bigvee_{q,q' \in Q} conf_{(q,(inc,i),q')}) \end{cases}$$

The next formula φ_{dec} is similar to φ_{inc}, the only difference being that a faulty behaviour cannot suppress the effect of a decrement. Thus, a *conf*-interval corresponding to a configuration produced by a decrement transition (i being the involved counter) is met by a *conf*-interval containing a c_i^--interval (recall that a c_i^+-interval occurring in a configuration encodes an increment transition *leading* to it, while a c_i^--interval occurring in a configuration encodes a decrement transition *starting* from it). In the next definition, $c_{i,b}^-$ plays the role played by $c_{i,b}^+$ in the definition of φ_{inc}.

$$\varphi_{dec} = \begin{cases} [G] \bigwedge_{i=1,\dots,|C|}(\langle \overline{A} \rangle c_{i,b}^- \leftrightarrow \langle \overline{A} \rangle c_i^-) \wedge \\ [G] \bigwedge_{i=1,\dots,|C|}(c_{i,b}^- \to (\langle \overline{A} \rangle conf \wedge [B] \neg conf)) \wedge \\ [G] \bigwedge_{i=1,\dots,|C|}(c_{i,b}^- \to [B] \neg c_{i,b}^-) \wedge \\ [G] \bigwedge_{i=1,\dots,|C|}((conf \wedge \langle \overline{A} \rangle (conf \wedge \langle B \rangle c_{i,b}^-)) \leftrightarrow \bigvee_{q,q' \in Q} conf_{(q,(dec,i),q')}) \end{cases}$$

The next formula φ_{ifz} states that a *conf*-interval corresponding to a configuration produced by a test transition (i being the involved counter) is met by a *conf*-interval that does not contain c_i-intervals.

$$\varphi_{ifz} = \begin{cases} [G] \bigwedge_{i=1,\dots,|C|}((\langle \overline{A} \rangle c_i \to [\overline{A}] c_{i,b}^z) \wedge (\neg \langle \overline{A} \rangle c_i \to [\overline{A}] \neg c_{i,b}^z)) \wedge \\ [G] \bigwedge_{(q,(ifz,i),q') \in \Delta}(conf_{(q,(ifz,i),q')} \to \langle \overline{A} \rangle (conf \wedge [\overline{B}] \neg c_{i,b}^z)) \end{cases}$$

Finally, the formula φ_{state} ensures that source and target states of transitions are correctly encoded by the relevant configurations:

$$\varphi_{state} = [G] \bigwedge_{(q,l,q') \in \Delta} ((conf_{(q,l,q')} \to \langle \overline{A} \rangle q) \wedge (\langle \overline{A} \rangle conf_{(q,l,q')} \to \langle \overline{A} \rangle q'))$$

Let $\varphi_{\mathcal{A}} = \varphi_{u-chain} \wedge \varphi_{units} \wedge \varphi_{conf-chain} \wedge \varphi_{corr} \wedge \varphi_{conf} \wedge \varphi_{inc} \wedge \varphi_{dec} \wedge \varphi_{ifz} \wedge \varphi_{state}$.

Lemma 14. *Let $M, [x, y] \Vdash \varphi_{\mathcal{A}}$. For every pair of consecutive conf-intervals $[x'_j, x'_{j+1}]$ and $[x'_{j+1}, x'_{j+2}]$, with $j \geq 0$, it holds that:*

- *(i) if $[x'_{j+1}, x'_{j+2}]$ contains a c_i^+-interval, then it is a $conf_{(q,(inc,i),q')}$-interval and it contains exactly one c_i^+-interval;*
- *(ii) $[x'_{j+1}, x'_{j+2}]$ is a $conf_{(q,(dec,i),q')}$-interval if and only $[x'_j, x'_{j+1}]$ contains a c_i^--interval; moreover, $[x'_j, x'_{j+1}]$ contains at most one c_i^--interval;*
- *(iii) if $[x'_{j+1}, x'_{j+2}]$ is a $conf_{(q,(ifz,i),q')}$-interval, then $[x'_j, x'_{j+1}]$ does not contain any c_i-interval;*
- *(iv) if $[x'_{j+1}, x'_{j+2}]$ is a $conf_{(q,l,q')}$-interval, then there exist x' and x'' such that $[x', x'_{j+1}]$ is a q-interval and $[x'', x'_{j+2}]$ is a q'-interval.*

Proof. The first two conjuncts of φ_{inc} force the proposition letter $c_{i,b}^+$ to hold exactly over intervals $[w, z]$ such that, for some $j > 0$, (a) $w = x'_j$, (b) $x'_j < z < x'_{j+1}$, and (c) there exists $w' \geq x'_j$ such that $[w', z]$ is a c_i^+-interval. This allows us to conclude that a *conf*-interval contains a c_i^+-interval if and only if among its prefixes there is a $c_{i,b}^+$-interval. More precisely, we can pair each c_i^+-interval contained in a *conf*-interval with a corresponding prefix $c_{i,b}^+$-interval. The third conjunct of φ_{inc} ensures that no *conf*-interval contains more than one $c_{i,b}^+$-interval as prefix, and the last one guarantees that every *conf*-interval containing a $c_{i,b}^+$-interval as a prefix is a $conf_{(q,(inc,i),q')}$-interval. Therefore, (i) holds.

As for (ii), the first two conjuncts of φ_{dec} force the proposition letter $c_{i,b}^-$ to behave (with respect to c_i^--intervals) exactly as $c_{i,b}^+$ does (with respect to c_i^+-intervals). Thus, the last conjunct forces every *conf*-interval met by a *conf*-interval that contains a c_i^--interval to be a $conf_{(q,(dec,i),q')}$-interval (left-to-right direction) and, conversely, every $conf_{(q,(dec,i),q')}$- interval to be met by a *conf*-interval containing a c_i^--interval (right-to-left direction). As in the previous case, the third conjunct ensures that no *conf*-interval contains more than one c_i^-- interval.

As for (iii), the first conjunct of φ_{ifz} forces the proposition letter $c_{i,b}^z$ to hold exactly over intervals $[w, z]$ such that z is the right endpoint of a c_i-interval. Hence, a *conf*-interval contains a c_i-interval if and only if it has a $c_{i,b}^z$-interval as a prefix. The second conjunct ensures that every $conf_{(q,(ifz,i),q')}$-interval is met by a *conf*-interval that does not contain any $c_{i,b}^z$-interval as a prefix.

Finally, (iv) immediately follows from φ_{state}. $\qquad\square$

The next lemma proves the correctness of the proposed encoding.

Lemma 15. $\varphi_{\mathcal{A}}$ *is satisfiable if and only if $\mathcal{A}$ has at least one infinite run.*

Proof. As for the left-to-right direction, let $\varphi_{\mathcal{A}}$ be satisfiable, that is, there exist an interval model M and an interval $[x, y]$ such that $M, [x, y] \Vdash \varphi_{\mathcal{A}}$. By Lemmas 10–14, we can easily check that M represents an infinite run of $\mathcal{A}$.

As for the converse direction, let $(q^0, \bar{v}^0), (q^1, \bar{v}^1), (q^2, \bar{v}^2), \ldots$ be an infinite run of $\mathcal{A}$. A model over $\mathbb{Q}$ that encodes the run can be built as follows. Let $(q, \bar{v})$ be the j-th configuration of the run ($j \geq 0$). It is encoded by the *conf*-interval $[x'_j, x'_{j+1}] = [\sum_{k=0}^{j} \frac{1}{2^k}, \sum_{k=0}^{j+1} \frac{1}{2^k}]$. $[x'_j, x'_{j+1}]$ is in turn partitioned into $1 + \sum_{k=1}^{|C|} v_k$ u-intervals with the same length (the sum of all counters' values, plus an interval for the state—recall that v_k is the k-th component of the vector $\bar{v}$), as follows: the first v_1 intervals are c_1-intervals, the following v_2 intervals are c_2-intervals, and so on; the last interval is a q-interval (notice that density plays a role here in allowing one to divide a fixed-size interval into an arbitrary large number of intervals). It is easy to define a valuation for all the proposition letters according to their intended meaning, as described in the above encoding, and to obtain a model M such that $M, [2, 3] \Vdash \varphi_{\mathcal{A}}$ (notice that density is exploited again here to make it possible for an infinite sequence of *conf*-intervals to fit in a bounded subset, namely, $[1, 2[$, of $\mathbb{Q}$). $\quad\square$

A similar reduction can be provided for the fragment $\overline{\mathsf{A}}\mathsf{B}$.

Theorem 5. *The satisfiability problem for $\overline{\mathsf{A}}\mathsf{B}$ and $\overline{\mathsf{A}}\overline{\mathsf{B}}$ over $\mathbb{Q}$ is non-primitive recursive.*

8. Rational numbers, real numbers, and dense linear orders

In this short section, we briefly compare the behaviour of HS fragments over $\mathbb{Q}$, $\mathbb{R}$, and the class of (all) dense linear orders. We first show that, as pointed out in Section 1, their behaviour over $\mathbb{Q}$ is (essentially) the same as over the class of dense linear orders; then, we point out similarities and differences between the cases of $\mathbb{Q}$ and $\mathbb{R}$.

In order to show the substantial equivalence between the linear order of $\mathbb{Q}$ and the class of all dense linear orders, suppose that a formula φ of a given HS fragment is satisfied in some dense model (as a matter of fact, in many HS fragments density can be added as part of the formula). This amounts to say that the standard translation of φ in first-order logic, denoted by $ST(\varphi)$,

is satisfied in the respective first-order model with the linear ordering relation $<$ and extra binary relations for the proposition letters (not quantified over). The existence of such a model for the first-order logic formula $ST(\varphi)$ is equivalent to the existence of a model in some dense linear order of the respective existentially quantified second-order version of $ST(\varphi)$. Now, by the Downward Löwenheim-Skolem theorem, $ST(\varphi)$ has a countable model, the dense order of which, by Cantor's theorem, is isomorphic either to $\mathbb{Q}$, or to an interval on $\mathbb{Q}$, depending on the endpoints, if any. Using that isomorphism, one can re-build the entire model of $ST(\varphi)$ on $\mathbb{Q}$ or on an interval on $\mathbb{Q}$. As an immediate consequence, we have that any procedure, e.g., a tableau system, that allows one to decide whether a formula of a given HS fragment has a model over $\mathbb{Q}$ can be easily turned into a procedure to establish whether such a formula is satisfiable over the class of (all) dense linear order. All that is needed is to adjust the decision procedure to look for models over $\mathbb{Q}$ with possible endpoints.

As for the relationships between the behaviour of HS fragments over $\mathbb{Q}$ and over $\mathbb{R}$, on the one hand, it has been shown that their relative expressive power over the former and the latter is exactly the same [12]; on the other hand, it has been proved that decidability and complexity of their satisfiability problem may differ. More precisely, in [30], Della Monica et al. have shown that the set of satisfiable formulae of the fragments A and $\overline{\mathsf{A}}$ over $\mathbb{Q}$ is the same as the set of satisfiable formulae over $\mathbb{R}$, and thus the NExpTime decision procedure developed for the former can actually be used also for the latter, but this is not the case with the fragment $\mathsf{A}\overline{\mathsf{A}}$. As shown in [30], there exists an $\mathsf{A}\overline{\mathsf{A}}$ formula that separates $\mathbb{R}$ from $\mathbb{Q}$, and thus the decision procedure for $\mathsf{A}\overline{\mathsf{A}}$ over $\mathbb{Q}$ cannot be exploited to check $\mathsf{A}\overline{\mathsf{A}}$ satisfiability over $\mathbb{R}$. (As a matter of fact, the satisfiability problem for $\mathsf{A}\overline{\mathsf{A}}$ turns out to be NExpTime-complete both over $\mathbb{Q}$ [17] and over $\mathbb{R}$ [31].) Finally, while the fragments $\mathsf{A}\overline{\mathsf{A}}\mathsf{B}\overline{\mathsf{B}}$ and $\mathsf{A}\overline{\mathsf{A}}\mathsf{E}\overline{\mathsf{E}}$ are decidable over $\mathbb{Q}$ [15] (see Section 4), it has been shown that they are undecidable over $\mathbb{R}$ [32].

9. Conclusions

In this paper, we focused on the variety of fragments of Halpern and Shoham's interval temporal logic (HS) interpreted over the linear order of the rational numbers $\mathbb{Q}$ or, equivalently, the class of all dense linear orders. As a matter of fact, the relative expressive power of all HS fragments over $\mathbb{Q}$ has been studied in [12], where it has been shown that there exist precisely

996 expressively different fragments. Moreover, it was known from recent results that the satisfiability problem for 130 of them is decidable.

The main result of the present paper is a complete classification of the decidable fragments of HS, interpreted over $\mathbb{Q}$, by their computational complexity. The emerging picture turns out to be multifaceted, including NP-, PSPACE-, NEXPTIME-, EXPSPACE-complete, and even non-primitive recursive fragments. This work adds up to a number of recent results aimed at providing a complete classification of HS fragments that include, so far, the relative expressive power of the fragments, when interpreted over the class of all linear orders [12], and their classification by computational complexity with respect to the class of (strongly) discrete linear orders, the class of finite linear orders, and the linear orders of the natural numbers $\mathbb{N}$ and of the integers $\mathbb{Z}$ [11].

As pointed out in Section 8, while the relative expressive power of HS fragments over $\mathbb{Q}$ and $\mathbb{R}$ is the same, this is not the case with their computational complexities. Some HS fragments have the same computational complexity over $\mathbb{Q}$ and $\mathbb{R}$, while other ones are decidable over $\mathbb{Q}$ and undecidable over $\mathbb{R}$; moreover, the decidability/complexity of some HS fragments over $\mathbb{R}$ is still unknown. In general, while the hardness results we provided for $\mathbb{Q}$ can be easily replicated for $\mathbb{R}$, as they are all based on the construction of the *computation history* of some Turing Machine and they all involve a sort of discretization of the underlying linear order, the portability of the upper bounds we obtained for $\mathbb{Q}$ to $\mathbb{R}$ is definitely not guaranteed. Among the missing elements of the picture for $\mathbb{R}$, we are quite interested in the status of the satisfiability problem for $\mathsf{B\overline{B}D\overline{D}L\overline{L}}$ and its fragments. One of the reasons is that there seems to be a natural connection between the fragment $\mathsf{D\overline{D}L\overline{L}}$, interpreted over $\mathbb{R}$, and the temporal logic of the two dimensional Minkowski spacetime, which has been recently investigated by Hirsh and Reynolds [33].

Acknowledgements

55

References

References

[1] D. Bresolin, D. Della Monica, A. Montanari, P. Sala, G. Sciavicco, On the complexity of fragments of the modal logic of Allen's relations over dense structures, in: Proc. of the 9th International Conference on Language and Automata Theory and Applications (LATA), Vol. 8977 of LNCS, Springer, 2015, pp. 511–523.

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

[3] A. Molinari, A. Montanari, A. Murano, G. Perelli, A. Peron, Checking interval properties of computations, Acta Informatica 53 (6-8) (2016) 587–619.

[4] Z. Chaochen, M. R. Hansen, Duration calculus: A formal approach to real-time systems, EATCS: Monographs in Theoretical Computer Science, Springer, 2004.

[5] A. Lomuscio, J. Michaliszyn, Model checking multi-agent systems against epistemic HS specifications with regular expressions, in: Proc. of the 15th International Conference on Principles of Knowledge Representation and Reasoning (KR), AAAI Press, 2016, pp. 298–308.

[6] R. Gennari, S. Tonelli, P. Vittorini, An AI-based process for generating games from flat stories, in: Proc. of the 33rd SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence, 2013, pp. 337–350.

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

[8] S. Laban, A. El-Desouky, RISMA: A rule-based interval state machine algorithm for alerts generation, performance analysis and monitoring real-time data processing, in: Proc. of the European Geosciences Union General Assembly, Vol. 15 of Geophysical Research Abstracts, 2013, p. 3309.

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

[10] D. Bresolin, D. Della Monica, V. Goranko, A. Montanari, G. Sciavicco, The dark side of interval temporal logic: marking the undecidability border, Annals of Mathematics and Artificial Intelligence 71 (1-3) (2014) 41–83.

[11] D. Bresolin, D. Della Monica, A. Montanari, P. Sala, G. Sciavicco, Interval temporal logics over strongly discrete linear orders: Expressiveness and complexity, Theoretical Computer Science 560 (3) (2014) 269–291.

[12] L. Aceto, D. Della Monica, V. Goranko, A. Ingólfsdóttir, A. Montanari, G. Sciavicco, A complete classification of the expressiveness of interval logics of Allen's relations: The general and the dense cases, Acta Informatica 53 (3) (2016) 207–246.

[13] L. Aceto, D. Della Monica, A. Ingólfsdóttir, A. Montanari, G. Sciavicco, On the expressiveness of the interval logic of Allen's relations over finite and discrete linear orders, in: Proc. of the 14th European Conference on Logics in Artificial Intelligence (JELIA), Vol. 8761 of LNCS, Springer, 2014, pp. 267–281.

[14] D. Bresolin, D. Della Monica, V. Goranko, A. Montanari, G. Sciavicco, Decidable and undecidable fragments of Halpern and Shoham's interval temporal logic: Towards a complete classification, in: Proc. of the 15th International Conference on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR), Vol. 5330 of LNCS, Springer, 2008, pp. 590–604.

[15] A. Montanari, G. Puppis, P. Sala, Decidability of the interval temporal logic $A\overline{A}B\overline{B}$ over the rationals, in: Proc. of the 39th International Symposium on Mathematical Foundations of Computer Science (MFCS), Vol. 8634 of LNCS, Springer, 2014, pp. 451–463.

[16] D. Bresolin, A. Montanari, P. Sala, G. Sciavicco, What's decidable about Halpern and Shoham's interval logic? The maximal fragment $AB\overline{BL}$, in: Proc. of the 26th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), 2011, pp. 387–396.

[17] D. Bresolin, A. Montanari, P. Sala, G. Sciavicco, Optimal tableau systems for propositional neighborhood logic over all, dense, and discrete linear orders, in: Proc. of the 20th International Conference on

Automated Reasoning with Analytic Tableaux and Related Methods (TABLEAUX), Vol. 6793 of LNAI, Springer, 2011, pp. 73–87.

[18] I. Shapirovsky, On PSPACE-decidability in transitive modal logic, in: Proc. of the 5th Conference on Advances in Modal Logic (AiML), King's College Publications, 2004, pp. 269–287.

[19] D. Bresolin, V. Goranko, A. Montanari, P. Sala, Tableaux for logics of subinterval structures over dense orderings, Journal of Logic and Computation 20 (1) (2010) 133–166.

[20] A. Montanari, G. Puppis, P. Sala, A decidable weakening of compass logic based on cone-shaped cardinal directions, Logical Methods in Computer Science 11 (4) (2015) 1–32.

[21] Divvy data challenge, `http://www.divvybikes.com/datachallenge`, last checked on February 21st, 2019.

[22] A. Tansel, J. Clifford, S. K. Gadia, S. Jajodia, A. Segev, R. T. Snodgrass (Eds.), Temporal Databases: Theory, Design, and Implementation, Benjamin/Cummings, 1993.

[23] P. Terenziani, R. T. Snodgrass, Reconciling point-based and interval-based semantics in temporal relational databases: A treatment of the telic/atelic distinction, IEEE Transactions on Knowledge and Data Engineering 16 (5) (2004) 540–551.

[24] L. Bozzelli, A. Molinari, A. Montanari, A. Peron, P. Sala, Interval vs. Point Temporal Logic Model Checking: an Expressiveness Comparison, in: Proc. of the 36th Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS), 2016, pp. 26:1–26:14.

[25] P. Roeper, Intervals and tenses, Journal of Philosophical Logic 9 (1980) 451–469.

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

[27] Y. Venema, Expressiveness and completeness of an interval tense logic, Notre Dame Journal of Formal Logic 31 (4) (1990) 529–547.

[28] P. Van Emde Boas, The convenience of tilings, in: Complexity, Logic and Recursion Theory, Vol. 187 of Lecture Notes in Pure and Applied Mathematics, Marcel Dekker Inc., 1997, pp. 331–363.

[29] P. Schnoebelen, Lossy counter machines decidability cheat sheet, in: Proc. of the 4th International Workshop on Reachability Problems (RP), 2010, pp. 51–75.

[30] D. Della Monica, A. Montanari, P. Sala, The importance of the past in interval temporal logics: the case of Propositional Neighborhood Logic, in: Logic Programs, Norms and Action - Essays in Honor of Marek J. Sergot on the Occasion of His 60th Birthday, Vol. 7360 of LNCS, Springer, 2012, pp. 79–102.

[31] A. Montanari, P. Sala, 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), 2012, pp. 39–46.

[32] A. Montanari, G. Puppis, P. Sala, 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), Vol. 6199 of LNCS, Springer, 2010, pp. 345–356.

[33] R. Hirsch, M. Reynolds, The temporal logic of two dimensional Minkowski spacetime is decidable, Journal of Symbolic Logic 83 (3) (2018) 829–867.