University of Udine

Department of Mathematics and Computer Science

PREPRINT

Interval vs. Point Temporal Logic Model Checking: an Expressiveness Comparison

Laura Bozzelli, Alberto Molinari, Angelo Montanari, Adriano Peron, and Pietro Sala

Preprint nr.: 3/2016

Reports available from: https://www.dimi.uniud.it/preprints/

Interval vs. Point Temporal Logic Model Checking: an Expressiveness Comparison

Laura Bozzelli[1], Alberto Molinari[2], Angelo Montanari[2], Adriano Peron[3], and Pietro Sala[4]

1 **Technical University of Madrid (UPM), Madrid, Spain**
2 **University of Udine, Udine, Italy**
3 **University of Napoli "Federico II", Napoli, Italy**
4 **University of Verona, Verona, Italy**

Abstract

Model checking is a powerful method widely explored in formal verification to check the (state-transition) model of a system against desired properties of its behaviour. Classically, properties are expressed by formulas of a temporal logic, such as LTL, CTL, and CTL*. These logics are "point-wise" interpreted, as they describe how the system evolves state-by-state. On the contrary, Halpern and Shoham's interval temporal logic (HS) is "interval-wise" interpreted, thus allowing one to naturally express properties of computation stretches, spanning a sequence of states, or properties involving temporal aggregations, which are inherently "interval-based".

In this paper, we study the expressiveness of HS in model checking, in comparison with that of the standard logics LTL, CTL, and CTL*. To this end, we consider HS endowed with three semantic variants: the state-based semantics, introduced by Montanari et al., which allows branching in the past and in the future, the linear-past semantics, allowing branching only in the future, and the linear semantics, disallowing branching. These variants are compared, as for their expressiveness, among themselves and to standard temporal logics, getting a complete picture. In particular, HS with linear (resp., linear-past) semantics is proved to be equivalent to LTL (resp., finitary CTL*).

1998 ACM Subject Classification F.4.1 Mathematical Logic, D.2.4 Software Verification

Keywords and phrases Interval Temporal Logics, Expressiveness, Model Checking

Digital Object Identifier 10.4230/LIPIcs...

1 Introduction

Point-based temporal logics (PTLs) provide a fundamental framework for the specification of the behavior of reactive systems, that makes it possible to describe how the system evolves state-by-state ("point-wise" view). PTLs have been successfully employed in *model checking* (MC), which enables one to automatically verify complex finite-state systems usually modelled as finite propositional Kripke structures. The MC methodology considers two types of PTLs—*linear* and *branching*—which differ in the underlying model of time. In linear temporal logics, such as LTL [24], each moment in time has a unique possible future: formulas are interpreted over paths of a Kripke structure, and thus they refer to a single computation of the system. In branching temporal logics, such as CTL and CTL* [7], each moment in time may evolve into several possible futures: formulas are interpreted over states of the Kripke structure, hence referring to all the possible computations of a system.

Interval temporal logics (ITLs) have been proposed as an alternative setting for reasoning about time [9, 23, 28]. Unlike standard PTLs, they take intervals, rather than points, as their primitive entities. ITLs allow one to specify relevant temporal properties that involve, for instance, actions with duration, accomplishments, and temporal aggregations, which are inherently "interval-based", and thus cannot be naturally expressed by PTLs. They have been applied in various areas of computer science, including formal verification, computational linguistics, planning, and multi-agent systems [14, 23, 25]. Halpern and Shoham's modal logic of time intervals HS [9] is the most popular among ITLs. It features one modality for each of the 13 possible ordering relations between pairs of intervals (the so-called Allen's relations [1]), apart from equality. Its *satisfiability problem* turns out to be undecidable for all interesting (classes of) linear orders [9]; the same happens with most of its fragments [6, 13, 17].

In this paper, we focus on the *model checking problem* for HS. In order to check interval properties of computations, one needs to collect information about states into computation stretches (i.e., finite paths of the Kripke structure, *tracks* for short): each track is interpreted as an interval, whose labelling is defined on the basis of the labelling of the component states. This approach to MC has independently and simultaneously been proposed by Montanari et al. in [22] and by Lomuscio and Michaliszyn in [14, 15].

The semantics proposed in [22] is *state-based*, featuring intervals/tracks which are forgetful of the history leading to the starting state of the interval itself. Since the starting state (resp., ending state) of an interval may feature several predecessors (resp., successors), this interpretation induces a branching reference in both future and past. The other relevant choice in this approach concerns the labeling of intervals: a natural principle, known as the *homogeneity assumption*, is adopted, according to which a proposition holds over an interval if and only if it holds over each component state. Under this semantics, the MC problem for full HS turns out to be decidable—it is **EXPSPACE**-hard, while the only known upper bound is non-elementary. The exact complexity of almost all the meaningful syntactic fragments of HS has been recently determined in a series of papers (e.g., [4, 5, 18, 19, 20, 21]).

The approach followed in [14, 15] is more expressive than the one in [22] since it relies on the extension of HS with knowledge modalities typical of the epistemic logics, which allow one to relate distinct paths of a Kripke structure. Additionally, the semantic assumptions differ from those of [22]: the logic is interpreted over the unwinding of the Kripke structure (*computation-tree-based* approach), and the interval labeling takes into account only the endpoints of the interval itself. A more expressive definition of interval labeling, obtained by associating each proposition with a regular expression over the set of states of the Kripke structure, was recently proposed in [16]. The decidability status of MC for full epistemic HS is currently unknown [14, 15].

In this paper, we study the expressiveness of HS, in the context of MC, in comparison with that of the standard PTLs LTL, CTL, and CTL*. The investigation is carried on enforcing the homogeneity assumption. We prove that HS endowed with the state-based semantics proposed in [22] (hereafter denoted as $\mathsf{HS}_{\mathsf{st}}$) is not comparable with LTL, CTL, and CTL*. On the one hand, the result supports the intuition that $\mathsf{HS}_{\mathsf{st}}$ gains some expressiveness by the ability to branch in the past. On the other hand, $\mathsf{HS}_{\mathsf{st}}$ does not feature the possibility to force the verification of a property over an infinite path, thus implying that the formalisms are not comparable. With the aim of having a more "effective" comparison base, we consider two semantic variants of HS, besides the state-based semantics $\mathsf{HS}_{\mathsf{st}}$, namely, the *computation-tree-based semantics* (denoted as $\mathsf{HS}_{\mathsf{lp}}$) and the *trace-based semantics* ($\mathsf{HS}_{\mathsf{lin}}$).

The state-based and computation-tree-based approaches rely on a branching-time setting and differ in the nature of past. In the latter approach, the past is linear: each interval may have several possible futures, but it has a unique past. Moreover, the past is assumed to be finite and cumulative (i.e., the history of the current situation increases with time, and is never forgotten). The trace-based approach relies on a linear-time setting, where the infinite paths (computations) of the given Kripke structure are the main semantic entities. Branching is neither allowed in the past nor in the future.

The variant $\mathsf{HS}_{\mathsf{lp}}$ is a natural candidate for an expressiveness comparison with the branching time logics CTL and CTL*. The more interesting and technically involved result is the characterization of $\mathsf{HS}_{\mathsf{lp}}$, which turns out to be expressively equivalent to finitary CTL*, i.e., the variant of CTL* with quantification over finite paths. As for CTL, a non comparability result can be stated. Conversely, $\mathsf{HS}_{\mathsf{lin}}$ is a natural candidate for an expressiveness comparison with LTL. As a matter of fact, we prove that $\mathsf{HS}_{\mathsf{lin}}$ and LTL are equivalent (even for a small fragment of $\mathsf{HS}_{\mathsf{lin}}$). We complete the picture with a comparison of the three semantic variants $\mathsf{HS}_{\mathsf{st}}$, $\mathsf{HS}_{\mathsf{lp}}$, and $\mathsf{HS}_{\mathsf{lin}}$. We prove that, as expected, $\mathsf{HS}_{\mathsf{lin}}$ is not comparable with either the branching versions, $\mathsf{HS}_{\mathsf{lp}}$ and $\mathsf{HS}_{\mathsf{st}}$. The interesting result is that, on the other hand, $\mathsf{HS}_{\mathsf{lp}}$ is strictly included in $\mathsf{HS}_{\mathsf{st}}$: this supports $\mathsf{HS}_{\mathsf{st}}$, adopted in [18, 19, 20, 21, 4, 5], as a reasonable and adequate semantic choice. The complete picture of the expressiveness results is reported in Figure 1 (the symbols $\neq$, $\equiv$ and $<$ denote incomparability, equivalence, and strict expressiveness inclusion, respectively).

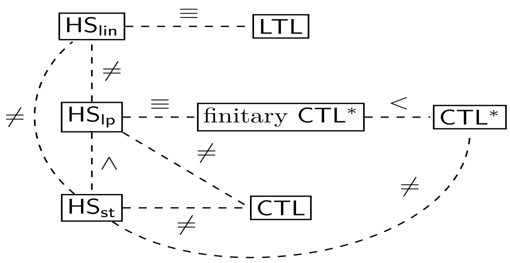


Figure 1 Overview of the expressiveness results

The paper is structured as follows. In Section 2, we introduce some preliminary notions. In Section 3, we prove the expressiveness results. In particular, in Section 3.1 we prove the equivalence between LTL and $\mathsf{HS}_{\mathsf{lin}}$; in Section 3.2 we prove the equivalence between $\mathsf{HS}_{\mathsf{lp}}$ and finitary CTL*; finally, in Section 3.3 we compare the logics $\mathsf{HS}_{\mathsf{st}}$, $\mathsf{HS}_{\mathsf{lp}}$, and $\mathsf{HS}_{\mathsf{lin}}$.

2 Preliminaries

Let $(\mathbb{N}, <)$ be the set of natural numbers equipped with the standard linear ordering. For all $i, j \in \mathbb{N}$, with $i \leq j$, $[i, j]$ denotes the set of natural numbers h such that $i \leq h \leq j$.

Let Σ be an alphabet and w be a non-empty finite or infinite word over Σ. We denote by $|w|$ the length of w (we set $|w| = \infty$ if w is infinite). For all $i, j \in \mathbb{N}$, with $i \leq j$, $w(i)$ denotes the i-th letter of w, while $w[i, j]$ denotes the finite subword of w given by $w(i) \cdots w(j)$. If w is finite and $|w| = n + 1$, we define $\mathrm{fst}(w) = w(0)$ and $\mathrm{lst}(w) = w(n)$. $\mathrm{Pref}(w) = \{w[0, i] \mid 0 \leq i \leq n - 1\}$ and $\mathrm{Suff}(w) = \{w[i, n] \mid 1 \leq i \leq n\}$ are the sets of all proper prefixes and suffixes of w, respectively.

2.1 Kripke structures and interval structures

▶ **Definition 1** (Kripke structure). A *Kripke structure* over a finite set $\mathcal{AP}$ of proposition letters is a tuple $\mathcal{K} = (\mathcal{AP}, S, \delta, \mu, s_0)$, where S is a set of states, $\delta \subseteq S \times S$ is a left-total transition relation, $\mu : S \mapsto 2^{\mathcal{AP}}$ is a total labelling function assigning to each state s the set of propositions that hold over it, and $s_0 \in S$ is the initial state. For $(s, s') \in \delta$, we say that s' is a successor of s, and s is a predecessor of s'. Finally, we say that $\mathcal{K}$ is finite if S is finite.

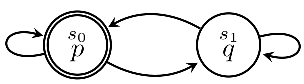


Figure 2 The Kripke structure $\mathcal{K}$

Figure 2 depicts the finite Kripke structure $\mathcal{K} = (\{p, q\}, \{s_0, s_1\}, \delta, \mu, s_0)$, where $\delta = \{(s_i, s_j) \mid i, j = 0, 1\}$, $\mu(s_0) = \{p\}$, and $\mu(s_1) = \{q\}$. The initial state s_0 is marked by a double circle.

Let $\mathcal{K} = (\mathcal{AP}, S, \delta, \mu, s_0)$ be a Kripke structure. An infinite path π of $\mathcal{K}$ is an infinite word over S such that $(\pi(i), \pi(i+1)) \in \delta$ for all $i \geq 0$. A *track* (or finite path) of $\mathcal{K}$ is a non-empty prefix of some infinite path of $\mathcal{K}$. A finite or infinite path is *initial* if it starts from the initial state of $\mathcal{K}$. Let $\mathrm{Trk}_{\mathcal{K}}$ be the (*infinite*) set of all tracks of $\mathcal{K}$ and $\mathrm{Trk}^0_{\mathcal{K}}$ be the set of initial tracks of $\mathcal{K}$. For a track ρ, $\mathrm{states}(\rho)$ denotes the set of states occurring in ρ, i.e., $\mathrm{states}(\rho) = \{\rho(0), \cdots, \rho(n)\}$, where $|\rho| = n + 1$.

▶ **Definition 2** (D-tree structure)**.** For a given set D of directions, a D*-tree structure* (over $\mathcal{AP}$) is a Kripke structure $\mathcal{K} = (\mathcal{AP}, S, \delta, \mu, s_0)$ such that $s_0 \in D$, S is a prefix closed subset of D^+, and δ is the set of pairs $(s, s') \in S \times S$ such that there exists $d \in D$ for which $s' = s \cdot d$ (note that δ is completely specified by S). The states of a D-tree structure are called *nodes*.

A Kripke structure $\mathcal{K} = (\mathcal{AP}, S, \delta, \mu, s_0)$ induces an S-tree structure, called the *computation tree of* $\mathcal{K}$, denoted by $\mathcal{C}(\mathcal{K})$, which is obtained by unwinding $\mathcal{K}$ from the initial state. Formally, $\mathcal{C}(\mathcal{K}) = (\mathcal{AP}, \mathrm{Trk}^0_{\mathcal{K}}, \delta', \mu', s_0)$, where the set of nodes is the set of initial tracks of $\mathcal{K}$ and for all $\rho, \rho' \in \mathrm{Trk}^0_{\mathcal{K}}$, $\mu'(\rho) = \mu(\mathrm{lst}(\rho))$ and $(\rho, \rho') \in \delta'$ iff $\rho' = \rho \cdot s$ for some $s \in S$.

Given a strict partial ordering $\mathbb{S} = (X, <)$, an *interval* in $\mathbb{S}$ is an ordered pair $[x, y]$ such that $x, y \in X$ and $x \leq y$. The interval $[x, y]$ denotes the subset of X given by the set of points $z \in X$ such that $x \leq z \leq y$. We denote by $\mathbb{I}(\mathbb{S})$ the set of intervals in $\mathbb{S}$.

▶ **Definition 3** (Interval structure)**.** An *interval structure* $\mathcal{IS}$ over $\mathcal{AP}$ is a pair $\mathcal{IS} = (\mathbb{S}, \sigma)$ such that $\mathbb{S} = (X, <)$ is a strict partial ordering and $\sigma : \mathbb{I}(\mathbb{S}) \mapsto 2^{\mathcal{AP}}$ is a labeling function assigning a set of proposition letters to each interval over $\mathbb{S}$.

2.2 Standard temporal logics

In this subsection, we recall the standard propositional temporal logics CTL^*, CTL, and LTL [7, 24]. For a set of proposition letters $\mathcal{AP}$, the formulas φ of CTL^* are defined as follows:

$$\varphi ::= \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \mathsf{X}\varphi \mid \varphi \mathsf{U} \varphi \mid \exists\varphi,$$

where $p \in \mathcal{AP}$, X and U are the "next" and "until" temporal modalities, and $\exists$ is the existential path quantifier. We also use standard shorthands: $\forall\varphi := \neg\exists\neg\varphi$ ("universal path quantifier"), $\mathsf{F}\varphi := \top\mathsf{U}\varphi$ ("eventually") and its dual $\mathsf{G}\varphi := \neg\mathsf{F}\neg\varphi$ ("always"). The logic CTL is the fragment of CTL^* where each temporal modality is immediately preceded by a path quantifier, while LTL corresponds to the fragment of the formulas devoid of path quantifiers.

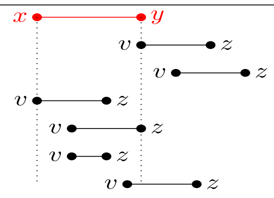
Given a Kripke structure $\mathcal{K} = (\mathcal{AP}, S, \delta, \mu, s_0)$, an infinite path π of $\mathcal{K}$, and a position $i \geq 0$ along π, the satisfaction relation $\mathcal{K}, \pi, i \models \varphi$ for CTL^*, written simply $\pi, i \models \varphi$ when $\mathcal{K}$ is clear from the context, is defined as follows (Boolean connectives are treated as usual):

$$\begin{array}{ll}
\pi, i \models p & \Leftrightarrow p \in \mu(\pi(i)), \\
\pi, i \models \mathsf{X}\varphi & \Leftrightarrow \pi, i+1 \models \varphi, \\
\pi, i \models \varphi_1 \mathsf{U} \varphi_2 & \Leftrightarrow \text{for some } j \geq i : \pi, j \models \varphi_2 \text{ and } \pi, k \models \varphi_1 \text{ for all } i \leq k < j, \\
\pi, i \models \exists\varphi & \Leftrightarrow \text{for some infinite path } \pi' \text{ starting from } \pi(i),\ \pi', 0 \models \varphi.
\end{array}$$

We say that $\mathcal{K}$ is a model of φ, written $\mathcal{K} \models \varphi$, if for all initial infinite paths π of $\mathcal{K}$, it holds that $\mathcal{K}, \pi, 0 \models \varphi$. We also consider a variant of CTL^*, called *finitary* CTL^*, where the path quantifier $\exists$ of CTL^* is replaced with the finitary path quantifier $\exists_f$. In this setting, path quantification ranges over the tracks (finite paths) starting from the current state. The satisfaction relation $\rho, i \models \varphi$, where ρ is a track and i is a position along ρ, is similar to that given for CTL^* with the only difference of finiteness of paths, and the fact that for a formula $\mathsf{X}\varphi$, $\rho, i \models \mathsf{X}\varphi$ iff $i + 1 < |\rho|$ and $\rho, i + 1 \models \varphi$. A Kripke structure $\mathcal{K}$ is a model of a finitary CTL^* formula if for each initial track ρ of $\mathcal{K}$, it holds that $\mathcal{K}, \rho, 0 \models \varphi$.

Table 1 Allen's relations and corresponding HS modalities

Allen relation	HS	Definition w.r.t. interval structures	Example
MEETS	$\langle A \rangle$	$[x,y]\mathcal{R}_A[v,z] \iff y = v$	
BEFORE	$\langle L \rangle$	$[x,y]\mathcal{R}_L[v,z] \iff y < v$	
STARTED-BY	$\langle B \rangle$	$[x,y]\mathcal{R}_B[v,z] \iff x = v \wedge z < y$	
FINISHED-BY	$\langle E \rangle$	$[x,y]\mathcal{R}_E[v,z] \iff y = z \wedge x < v$	
CONTAINS	$\langle D \rangle$	$[x,y]\mathcal{R}_D[v,z] \iff x < v \wedge z < y$	
OVERLAPS	$\langle O \rangle$	$[x,y]\mathcal{R}_O[v,z] \iff x < v < y < z$	



2.3 The interval temporal logic HS

An interval algebra was proposed by Allen in [1] to reason about intervals and their relative order, while a systematic logical study of interval representation and reasoning was done a few years later by Halpern and Shoham, who introduced the interval temporal logic HS featuring one modality for each Allen relation, but equality [9]. Table 1 depicts 6 of the 13 Allen's relations, together with the corresponding HS (existential) modalities. The other 7 relations are the 6 inverse relations (given a binary relation $\mathcal{R}$, the inverse relation $\overline{\mathcal{R}}$ is such that $b\overline{\mathcal{R}}a$ if and only if $a\mathcal{R}b$) and equality.

For a set of proposition letters $\mathcal{AP}$, the formulas ψ of HS are defined as follows:

$$\psi ::= p \mid \neg\psi \mid \psi \wedge \psi \mid \langle X \rangle \psi,$$

where $p \in \mathcal{AP}$ and $X \in \{A, L, B, E, D, O, \overline{A}, \overline{L}, \overline{B}, \overline{E}, \overline{D}, \overline{O}\}$. For any modality $\langle X \rangle$, the dual universal modality $[X]\psi$ is defined as $\neg\langle X \rangle\neg\psi$. For any subset of Allen's relations $\{X_1, \ldots, X_n\}$, let $\mathsf{X_1 \cdots X_n}$ be the HS fragment featuring modalities for $X_1, \ldots, X_n$ only.

We assume the *non-strict semantics of HS*, which admits intervals consisting of a single point.[1] Under such an assumption, all HS modalities can be expressed in terms of modalities $\langle B \rangle, \langle E \rangle, \langle \overline{B} \rangle$, and $\langle \overline{E} \rangle$ [28], e.g., modality $\langle A \rangle$ can be expressed in terms of $\langle E \rangle$ and $\langle \overline{B} \rangle$ as $\langle A \rangle\, \varphi := ([E]\bot \wedge (\varphi \vee \langle \overline{B} \rangle\, \varphi)) \vee \langle E \rangle([E]\bot \wedge (\varphi \vee \langle \overline{B} \rangle\, \varphi))$. We also use the derived operator $\langle G \rangle$ of HS (and its dual $[G]$), which allows one to select arbitrary subintervals of the given interval and is defined as: $\langle G \rangle\, \psi := \psi \vee \langle B \rangle\, \psi \vee \langle E \rangle\, \psi \vee \langle B \rangle\, \langle E \rangle\, \psi$. HS can be viewed as a multi-modal logic with $\langle B \rangle, \langle E \rangle, \langle \overline{B} \rangle$, and $\langle \overline{E} \rangle$ as primitive modalities and its semantics can be defined over a multi-modal Kripke structure, called *abstract interval model*, where intervals are treated as atomic objects and Allen's relations as binary relations over intervals.

▶ **Definition 4** (Abstract interval model [18])**.** An *abstract interval model* over $\mathcal{AP}$ is a tuple $\mathcal{A} = (\mathcal{AP}, \mathbb{I}, B_\mathbb{I}, E_\mathbb{I}, \sigma)$, where $\mathbb{I}$ is a set of worlds, $B_\mathbb{I}$ and $E_\mathbb{I}$ are two binary relations over $\mathbb{I}$, and $\sigma : \mathbb{I} \mapsto 2^{\mathcal{AP}}$ is a labeling function assigning a set of proposition letters to each world.

Let $\mathcal{A} = (\mathcal{AP}, \mathbb{I}, B_\mathbb{I}, E_\mathbb{I}, \sigma)$ be an abstract interval model. In the interval setting, $\mathbb{I}$ is interpreted as a set of intervals, and $B_\mathbb{I}$ and $E_\mathbb{I}$ as Allen's relations B (*started-by*) and E (*finished-by*), respectively; σ assigns to each interval in $\mathbb{I}$ the set of proposition letters that hold over it. Given an interval $I \in \mathbb{I}$, the truth of an HS formula over I is inductively defined as follows (Boolean connectives are treated as usual):

- $\mathcal{A}, I \models p$ iff $p \in \sigma(I)$, for any $p \in \mathcal{AP}$;
- $\mathcal{A}, I \models \langle X \rangle\psi$, for $X \in \{B, E\}$, iff there exists $J \in \mathbb{I}$ such that $I\, X_\mathbb{I}\, J$ and $\mathcal{A}, J \models \psi$;
- $\mathcal{A}, I \models \langle \overline{X} \rangle\psi$, for $\overline{X} \in \{\overline{B}, \overline{E}\}$, iff there exists $J \in \mathbb{I}$ such that $J\, X_\mathbb{I}\, I$ and $\mathcal{A}, J \models \psi$.

[1] All the results we prove in the paper hold for the strict semantics as well.

▶ **Definition 5** (Abstract interval model induced by an interval structure). An interval structure $\mathcal{IS} = (\mathbb{S}, \sigma)$, with $\mathbb{S} = (X, <)$, *induces* the abstract interval model $\mathcal{A}_{\mathcal{IS}} = (\mathcal{AP}, \mathbb{I}(\mathbb{S}), B_{\mathbb{I}(\mathbb{S})}, E_{\mathbb{I}(\mathbb{S})}, \sigma)$, where $[x, y]\, B_{\mathbb{I}(\mathbb{S})}\, [v, z]$ iff $x = v$ and $z < y$, and $[x, y]\, E_{\mathbb{I}(\mathbb{S})}\, [v, z]$ iff $y = z$ and $x < v$.
For an interval I and an HS formula ψ, we write $\mathcal{IS}, I \models \psi$ to mean that $\mathcal{A}_{\mathcal{IS}}, I \models \psi$.

2.4 Three variants of HS semantics for model checking

In this section, we define the three variants of HS semantics $\mathsf{HS}_{\mathsf{st}}$ (state-based semantics), $\mathsf{HS}_{\mathsf{lp}}$ (computation-tree-based semantics), and $\mathsf{HS}_{\mathsf{lin}}$ (trace-based semantics) for model checking HS against Kripke structures. For each such variant $\mathcal{S}$, the related (finite) model checking problem is deciding whether a finite Kripke structure is a model of an HS formula under $\mathcal{S}$.

Let us start with the *state-based semantics* [22], where an abstract interval model is naturally associated with a given Kripke structure $\mathcal{K}$ by considering the set of intervals as the set $\mathrm{Trk}_{\mathcal{K}}$ of tracks of $\mathcal{K}$.

▶ **Definition 6** (Abstract interval model induced by a Kripke structure). The *abstract interval model induced by a Kripke structure* $\mathcal{K} = (\mathcal{AP}, S, \delta, \mu, s_0)$ is $\mathcal{A}_{\mathcal{K}} = (\mathcal{AP}, \mathbb{I}, B_{\mathbb{I}}, E_{\mathbb{I}}, \sigma)$, where $\mathbb{I} = \mathrm{Trk}_{\mathcal{K}}$, $B_{\mathbb{I}} = \{(\rho, \rho') \in \mathbb{I} \times \mathbb{I} \mid \rho' \in \mathrm{Pref}(\rho)\}$, $E_{\mathbb{I}} = \{(\rho, \rho') \in \mathbb{I} \times \mathbb{I} \mid \rho' \in \mathrm{Suff}(\rho)\}$, and $\sigma : \mathbb{I} \mapsto 2^{\mathcal{AP}}$ is such that $\sigma(\rho) = \bigcap_{s \in \mathrm{states}(\rho)} \mu(s)$, for all $\rho \in \mathbb{I}$.

According to the definition of σ, $p \in \mathcal{AP}$ holds over $\rho = v_1 \cdots v_n$ if and only if it holds over all the states $v_1, \ldots, v_n$ of ρ. This conforms to the *homogeneity principle*, according to which a proposition letter holds over an interval if and only if it holds over all its subintervals [26].

▶ **Definition 7** (State-based semantics). Let $\mathcal{K}$ be a Kripke structure and ψ be an HS formula. A track $\rho \in \mathrm{Trk}_{\mathcal{K}}$ satisfies ψ under the state-based semantics, denoted as $\mathcal{K}, \rho \models_{\mathsf{st}} \psi$, if it holds that $\mathcal{A}_{\mathcal{K}}, \rho \models \psi$. Moreover, $\mathcal{K}$ *is a model of* ψ *under the state-based semantics*, denoted as $\mathcal{K} \models_{\mathsf{st}} \psi$, if for all *initial* tracks $\rho \in \mathrm{Trk}^0_{\mathcal{K}}$, it holds that $\mathcal{K}, \rho \models_{\mathsf{st}} \psi$.

We now introduce the *computation-tree-based semantics*, where we simply consider the abstract interval model *induced by the computation tree* of the Kripke structure. Notice that since each state in a computation tree has a unique predecessor (with the exception of the initial state), this HS semantics induces a linear reference in the past.

▶ **Definition 8** (Computation-tree-based semantics). A Kripke structure $\mathcal{K}$ *is a model of an HS formula* ψ *under the computation-tree-based semantics*, written $\mathcal{K} \models_{\mathsf{lp}} \psi$, if $\mathcal{C}(\mathcal{K}) \models_{\mathsf{st}} \psi$.

Finally, we propose the *trace-based semantics*, which exploits the interval structures induced by the infinite paths of the Kripke structure.

▶ **Definition 9** (Interval structure induced by an infinite path). For a Kripke structure $\mathcal{K} = (\mathcal{AP}, S, \delta, \mu, s_0)$ and an infinite path $\pi = \pi(0)\pi(1) \cdots$ of $\mathcal{K}$, the *interval structure induced by* π is $\mathcal{IS}_{\mathcal{K},\pi} = ((\mathbb{N}, <), \sigma)$, where for each interval $[i, j]$, $\sigma([i, j]) = \bigcap_{h=i}^{j} \mu(\pi(h))$.

▶ **Definition 10** (Trace-based semantics). A Kripke structure $\mathcal{K}$ *is a model of an HS formula* ψ *under the trace-based semantics*, denoted as $\mathcal{K} \models_{\mathsf{lin}} \psi$, iff for each initial infinite path π and for each initial interval $[0, i]$, it holds that $\mathcal{IS}_{\mathcal{K},\pi}, [0, i] \models \psi$.

3 Expressiveness

In this section, we compare the expressive power of the logics $\mathsf{HS}_{\mathsf{st}}$, $\mathsf{HS}_{\mathsf{lp}}$, $\mathsf{HS}_{\mathsf{lin}}$, LTL, CTL, and CTL^* when interpreted over finite Kripke structures. Given two logics L_1 and L_2, and two formulas $\varphi_1 \in L_1$ and $\varphi_2 \in L_2$, we say that φ_1 in L_1 is *equivalent* to φ_2 in L_2 if, for every finite Kripke structure $\mathcal{K}$, $\mathcal{K}$ is a model of φ_1 in L_1 if and only if $\mathcal{K}$ is a model of φ_2 in L_2. When comparing the expressive power of two logics L_1 and L_2, we say that L_2 *is subsumed by* L_1, denoted as $L_1 \geq L_2$, if for each formula $\varphi_2 \in L_2$, there exists a formula $\varphi_1 \in L_1$ such that φ_1 in L_1 is equivalent to φ_2 in L_2. Moreover, L_1 is *as expressive as* L_2 (or, L_1 and L_2 have *the same expressiveness*), written $L_1 \equiv L_2$, if both $L_1 \geq L_2$ and $L_2 \geq L_1$. We say that L_1 is *more expressive than* L_2 if $L_1 \geq L_2$ and $L_2 \not\geq L_1$. Finally, L_1 and L_2 are *expressively incomparable* if both $L_1 \not\geq L_2$ and $L_2 \not\geq L_1$.

3.1 Equivalence between LTL and $\mathsf{HS}_{\mathsf{lin}}$

In this section we show that $\mathsf{HS}_{\mathsf{lin}}$ is as expressive as LTL even for small syntactical fragments of $\mathsf{HS}_{\mathsf{lin}}$. For this purpose, we exploit the well-known equivalence between LTL and First Order Logic (FO) over infinite words. Recall that given a countable set $\{x, y, z, \ldots\}$ of (position) variables, FO formulas φ over a set of proposition symbols $\mathcal{AP} = \{p, \ldots\}$ are defined as:

$$\varphi := \top \mid p \in x \mid x \leq y \mid x < y \mid \neg\,\varphi \mid \varphi \wedge \varphi \mid \exists x.\varphi\,.$$

We interpret FO formulas φ over infinite paths π of Kripke structures $\mathcal{K} = (\mathcal{AP}, S, \delta, \mu, s_0)$. Given a variable *valuation* g, assigning to each variable a position $i \geq 0$, the satisfaction relation $(\pi, g) \models \varphi$ corresponds to the standard satisfaction relation $(\mu(\pi), g) \models \varphi$, where $\mu(\pi)$ is the infinite word over $2^{\mathcal{AP}}$ given by $\mu(\pi(0))\mu(\pi(1))\cdots$ (for the details, see Appendix A). We write $\pi \models \varphi$ to mean that $(\pi, g_0) \models \varphi$, where $g_0(x) = 0$ for each variable x. An FO sentence is a formula with no free variables. The following is a well-known result [10].

▶ Proposition 1. Given a FO sentence φ over $\mathcal{AP}$, one can construct an LTL formula ψ such that for all Kripke structures $\mathcal{K}$ over $\mathcal{AP}$ and infinite paths π, it holds that $\pi \models \varphi$ iff $\pi, 0 \models \psi$.

Given a $\mathsf{HS}_{\mathsf{lin}}$ formula ψ, we construct a FO sentence ψ_{FO} such that, for all Kripke structures $\mathcal{K}$, $\mathcal{K} \models_{\mathsf{lin}} \psi$ iff for each initial infinite path π of $\mathcal{K}$, $\pi \models \psi_{\mathsf{FO}}$. The formula ψ_{FO} is given by $\exists x((\forall z.z \geq x) \wedge \forall y.h(\psi, x, y))$, where $h(\psi, x, y)$ is a FO formula having x and y as free variables (intuitively, representing the endpoints of the current interval) and ensuring that for each infinite path π and interval $[i, j]$, $\mathcal{IS}_{\mathcal{K},\pi}, [i, j] \models \psi$ iff $(\pi, g) \models h(\psi, x, y)$ for any valuation g such that $g(x) = i$ and $g(y) = j$. The construction of $h(\psi, x, y)$ is straightforward (for the details, see Appendix A). Thus, by Proposition 1, we obtain the following result.

▶ **Theorem 11.** *LTL* $\geq$ *$\mathsf{HS}_{\mathsf{lin}}$.*

Conversely, we show that LTL can be translated in linear-time into $\mathsf{HS}_{\mathsf{lin}}$ (actually, the fragment AB, featuring only modalities for A and B, is expressive enough for the purpose). In the following we will make use of the B formula length_n, with $n \geq 1$, characterizing the intervals of length n, which is defined as follows: $\mathit{length}_n := (\underbrace{\langle \mathrm{B}\rangle \ldots \langle \mathrm{B}\rangle}_{n-1} \top) \;\wedge\; (\underbrace{[B] \ldots [B]}_{n} \bot)$.

▶ **Theorem 12.** *Given an LTL formula φ, one can construct in linear-time an* AB *formula ψ such that φ in LTL is equivalent to ψ in* $\mathsf{AB}_{\mathsf{lin}}$.

Proof. Let $f : \mathsf{LTL} \mapsto \mathsf{AB}$ be the mapping homomorphic w. r. to the Boolean connectives, defined as follows for each proposition p and for the temporal modalities X and U:

$$f(p) = p, \qquad f(\mathsf{X}\psi) = \langle \mathrm{A}\rangle(\mathit{length}_2 \wedge \langle \mathrm{A}\rangle(\mathit{length}_1 \wedge f(\psi))),$$
$$f(\psi_1 \mathsf{U} \psi_2) = \langle \mathrm{A}\rangle\Big(\langle \mathrm{A}\rangle(\mathit{length}_1 \wedge f(\psi_2)) \wedge [B](\langle \mathrm{A}\rangle(\mathit{length}_1 \wedge f(\psi_1))\Big).$$

Given a Kripke structure $\mathcal{K}$, an infinite path π, a position $i \geq 0$, and an LTL formula ψ, by a straightforward induction on the structure of ψ we can show that $\pi, i \models \psi$ iff $\mathcal{IS}_{\mathcal{K},\pi}, [i,i] \models f(\psi)$. Hence $\mathcal{K} \models \psi$ iff $\mathcal{K} \models_{\mathsf{lin}} length_1 \rightarrow f(\psi)$. ◀

▶ **Corollary 13.** *$\mathsf{HS}_{\mathsf{lin}}$ and LTL have the same expressiveness.*

3.2 A characterization of $\mathsf{HS}_{\mathsf{lp}}$

In this section we show that $\mathsf{HS}_{\mathsf{lp}}$ is as expressive as *finitary* CTL*. Actually, the result can be proved to hold already for the syntactical fragment ABE (which does not feature transposed modalities). In addition, we show that $\mathsf{HS}_{\mathsf{lp}}$ is subsumed by CTL*.

We first show that finitary CTL* is subsumed by $\mathsf{HS}_{\mathsf{lp}}$. The result is proved by exploiting a preliminary property stating that, when interpreted over finite words, the BE fragment of HS and LTL define the same class of finitary languages. For an LTL formula φ with proposition symbols over an alphabet Σ (in our case Σ is $2^{\mathcal{AP}}$), $L_{act}(\varphi)$ denotes the set of non-empty finite words over Σ satisfying φ under the standard action-based semantics of LTL, interpreted over finite words (see [27]). A similar notion can be given for BE formulas φ with propositional symbols in Σ (considered under the homogeneity principle). Then φ denotes a language, written $L_{act}(\varphi)$, of non-empty finite words over Σ, inductively defined as:

- $L_{act}(a) = a^+$ for each $a \in \Sigma$;
- $L_{act}(\neg\varphi) = \Sigma^+ \setminus L_{act}(\varphi)$;
- $L_{act}(\varphi_1 \wedge \varphi_2) = L_{act}(\varphi_1) \cap L_{act}(\varphi_2)$;
- $L_{act}(\langle \mathrm{B} \rangle\, \varphi) = \{w \in \Sigma^+ \mid \mathrm{Pref}(w) \cap L_{act}(\varphi) \neq \emptyset\}$;
- $L_{act}(\langle \mathrm{E} \rangle\, \varphi) = \{w \in \Sigma^+ \mid \mathrm{Suff}(w) \cap L_{act}(\varphi) \neq \emptyset\}$.

We prove that under the action-based semantics, BE formulas and LTL formulas define the same class of finitary languages. By proceeding as in Section 3.1, one can easily show that, over finite words, the class of languages defined by the fragment BE is subsumed by that defined by LTL. To prove the converse direction we exploit an algebraic condition introduced in [29], here called *LTL-closure*, which gives, for a class of finitary languages, a sufficient condition to guarantee the inclusion of the class of LTL-definable languages.

▶ **Definition 14** (LTL-closure)**.** A class $\mathcal{C}$ of languages of finite words over finite alphabets is *LTL-closed* iff the following conditions are satisfied, where Σ and Δ are finite alphabets, $b \in \Sigma$ and $\Gamma = \Sigma \setminus \{b\}$:

1. $\mathcal{C}$ is closed under language complementation and language intersection.
2. If $L \in \mathcal{C}$ with $L \subseteq \Gamma^+$, then $\Sigma^* bL$, $\Sigma^* b(L+\varepsilon)$, $Lb\Sigma^*$, $(L+\varepsilon)b\Sigma^*$ are in $\mathcal{C}$.
3. Let $U_0 = \Gamma^* b$, $h_0 : U_0 \rightarrow \Delta$ and $h : U_0^+ \rightarrow \Delta^+$ be defined by $h(u_0 u_1 \dots u_n) = h_0(u_0) \dots h_0(u_n)$. Assume that for each $d \in \Delta$, the language $L_d = \{u \in \Gamma^+ \mid h_0(ub) = d\}$ is in $\mathcal{C}$. Then for each language $L \in \mathcal{C}$ s.t. $L \subseteq \Delta^+$, the language $\Gamma^* b h^{-1}(L) \Gamma^*$ is in $\mathcal{C}$.

▶ **Theorem 15** ([29])**.** *Any LTL-closed class $\mathcal{C}$ of finitary languages includes the class of LTL-definable finitary languages.*

▶ **Theorem 16.** *Let φ be an LTL formula over a finite alphabet Σ. Then there exists a* BE *formula φ_{HS} over Σ such that $L_{act}(\varphi_{\mathsf{HS}}) = L_{act}(\varphi)$.*

Proof. It suffices to prove that the class of finitary languages definable by BE formulas is LTL-closed, and to apply Theorem 15 (the proof of LTL-closure is reported in Appendix B). ◀

By exploiting Theorem 16, we establish the following result.

▶ **Theorem 17.** *Let φ be a finitary CTL* formula over $\mathcal{AP}$. Then there is an* ABE *formula φ_{HS} over $\mathcal{AP}$ s.t. for all Kripke structures $\mathcal{K}$ over $\mathcal{AP}$ and tracks ρ, $\mathcal{K}, \rho, 0 \models \varphi$ iff $\mathcal{K}, \rho \models_{\mathsf{st}} \varphi_{\mathsf{HS}}$.*

Proof. The proof is by induction on the nesting depth of modality $\exists_f$ in φ. The base case (φ is a finitary LTL formula over $\mathcal{AP}$) is similar to the inductive step, thus we can focus our attention on the latter. Let H be the non-empty set of subformulas of φ of the form $\exists_f \psi$ which do not occur in the scope of the path quantifier $\exists_f$. Then φ can be seen as an LTL formula over the set of atomic propositions $\overline{\mathcal{AP}} = \mathcal{AP} \cup H$. Let $\Sigma = 2^{\overline{\mathcal{AP}}}$ and $\overline{\varphi}$ be the LTL formula over Σ obtained from φ by replacing each occurrence of $p \in \overline{\mathcal{AP}}$ in φ with the formula $\bigvee_{P \in \Sigma \,:\, p \in P} P$, according to the LTL action-based semantics.

Given a Kripke structure $\mathcal{K}$ over $\mathcal{AP}$ with labeling μ and a track ρ of $\mathcal{K}$, we denote by ρ_H the finite word over $2^{\overline{\mathcal{AP}}}$ of length $|\rho|$ defined as $\rho_H(i) = \mu(\rho(i)) \cup \{\exists_f \psi \in H \mid \mathcal{K}, \rho, i \models \exists_f \psi\}$, for all $i \in [0, |\rho| - 1]$. One can easily show by structural induction on φ that

Claim 1: $\mathcal{K}, \rho, 0 \models \varphi$ iff $\rho_H \in L_{act}(\overline{\varphi})$.

By Theorem 16, there exists a BE formula $\overline{\varphi}_{\mathsf{HS}}$ over Σ such that $L_{act}(\overline{\varphi}) = L_{act}(\overline{\varphi}_{\mathsf{HS}})$. Moreover, by the induction hypothesis, for each formula $\exists_f \psi \in H$, there exists an ABE formula ψ_{HS} such that for all Kripke structures $\mathcal{K}$ and tracks ρ of $\mathcal{K}$, $\mathcal{K}, \rho, 0 \models \psi$ iff $\mathcal{K}, \rho \models_{\mathsf{st}} \psi_{\mathsf{HS}}$. Since ρ is arbitrary, $\mathcal{K}, \rho, i \models \exists_f \psi$ iff $\mathcal{K}, \rho[i,i], 0 \models \exists_f \psi$ iff $\mathcal{K}, \rho[i,i] \models_{\mathsf{st}} \langle \mathrm{A} \rangle\, \psi_{\mathsf{HS}}$, for each $i \geq 0$. Let φ_{HS} be the ABE formula over $\mathcal{AP}$ obtained from the BE formula $\overline{\varphi}_{\mathsf{HS}}$ by replacing each occurrence of $P \in \Sigma$ in $\overline{\varphi}_{\mathsf{HS}}$ with the formula

$$[G](\mathit{length}_1 \rightarrow \bigwedge_{\exists_f \psi \in H \cap P} \langle \mathrm{A} \rangle\, \psi_{\mathsf{HS}} \wedge \bigwedge_{\exists_f \psi \in H \setminus P} \neg \langle \mathrm{A} \rangle\, \psi_{\mathsf{HS}} \wedge \bigwedge_{p \in \mathcal{AP} \cap P} p \wedge \bigwedge_{p \in \mathcal{AP} \setminus P} \neg p).$$

Since for all $i \geq 0$ and $\exists_f \psi \in H$, $\mathcal{K}, \rho, i \models \exists_f \psi$ iff $\mathcal{K}, \rho[i,i] \models_{\mathsf{st}} \langle \mathrm{A} \rangle\, \psi_{\mathsf{HS}}$, by a straightforward induction on the structure of $\overline{\varphi}_{\mathsf{HS}}$, for all Kripke structures $\mathcal{K}$ and tracks ρ of $\mathcal{K}$ we have $\mathcal{K}, \rho \models_{\mathsf{st}} \varphi_{\mathsf{HS}}$ iff $\rho_H \in L_{act}(\overline{\varphi}_{\mathsf{HS}})$. Therefore, since $L_{act}(\overline{\varphi}) = L_{act}(\overline{\varphi}_{\mathsf{HS}})$, by Claim 1 $\mathcal{K}, \rho, 0 \models \varphi$ iff $\mathcal{K}, \rho \models_{\mathsf{st}} \varphi_{\mathsf{HS}}$, for arbitrary Kripke structures $\mathcal{K}$ and tracks ρ of $\mathcal{K}$. ◀

Since for the fragment ABE of HS the computation-tree-based semantics coincides with the state-based semantics, by Theorem 17 we obtain the following corollary.

▶ **Corollary 18.** *Finitary CTL* is subsumed by both HS$_{\mathsf{st}}$ and HS$_{\mathsf{lp}}$.*

Conversely, we show now that HS$_{\mathsf{lp}}$ is subsumed by both CTL* and the finitary variant of CTL*. For this purpose, we first introduce a hybrid and linear-past extension of CTL*, called *hybrid* CTL$^*_{lp}$, and its finitary variant, called *finitary hybrid* CTL$^*_{lp}$. Hybrid logics (see [3]), besides standard modalities, make use of explicit variables and quantifiers that bind them. Variables and binders allow us to easily mark points in a path, which will be considered as starting and ending points of intervals, thus permitting a natural encoding of HS$_{\mathsf{lp}}$. Actually, we will show that the restricted form of use of variables and binders exploited in our encoding does not increase the expressive power of (finitary) CTL* (as it happens for an unrestricted use), thus proving the desired result. We start by defining *hybrid* CTL$^*_{lp}$.

For a countable set $\{x, y, z, \ldots\}$ of (position) variables, the set of formulas φ of hybrid CTL$^*_{lp}$ over $\mathcal{AP}$ is defined as follows:

$$\varphi ::= \top \mid p \mid x \mid \neg\varphi \mid \varphi \vee \varphi \mid {\downarrow} x.\varphi \mid \mathsf{X}\varphi \mid \varphi \mathsf{U} \varphi \mid \mathsf{X}^{-}\varphi \mid \varphi \mathsf{U}^{-} \varphi \mid \exists\varphi,$$

where X^{-} ("previous") and U^{-} ("since") are the past counterparts of the "next" and "until" modalities X and U, and ${\downarrow} x$ is the *downarrow binder operator* [3], which binds x to the current position along the given initial infinite path. We also use the standard shorthands $\mathsf{F}^-\varphi := \top \mathsf{U}^- \varphi$ ("eventually in the past") and its dual $\mathsf{G}^-\varphi := \neg\mathsf{F}^-\neg\varphi$ ("always in the past"). As usual, a sentence is a formula with no free variables. Let $\mathcal{K}$ be a Kripke structure and φ be a hybrid CTL^*_{lp} formula. For an *initial* infinite path π of $\mathcal{K}$, a variable valuation g assigning to each variable x a position along π, and $i \geq 0$, the satisfaction relation $\pi, g, i \models \varphi$ is defined as follows (we omit the clauses for the Boolean connectives and for U and X):

$$\begin{array}{ll}
\pi, g, i \models \mathsf{X}^-\varphi & \Leftrightarrow i > 0 \text{ and } \pi, g, i-1 \models \varphi, \\
\pi, g, i \models \varphi_1 \mathsf{U}^- \varphi_2 & \Leftrightarrow \text{for some } j \leq i : \pi, g, j \models \varphi_2 \text{ and } \pi, g, k \models \varphi_1 \text{ for all } j < k \leq i, \\
\pi, g, i \models \exists\varphi & \Leftrightarrow \text{for some initial infinite path } \pi' \text{ s.t. } \pi'[0,i] = \pi[0,i],\ \pi', g, i \models \varphi, \\
\pi, g, i \models x & \Leftrightarrow g(x) = i, \\
\pi, g, i \models {\downarrow} x.\varphi & \Leftrightarrow \pi, g[x \leftarrow i], i \models \varphi,
\end{array}$$

where $g[x \leftarrow i](x) = i$ and $g[x \leftarrow i](y) = g(y)$ for $y \neq x$. A Kripke structure $\mathcal{K}$ is a model of a formula φ if for each initial infinite path π, $\pi, g_0, 0 \models \varphi$, where g_0 assigns 0 to each variable. Note that path quantification is "memoryful", i.e., it ranges over infinite paths that start at the root and visit the current node of the computation tree. Clearly, the semantics for the syntactical fragment CTL^* coincides with the standard one. If we discharge the use of variables and binder modalities, we obtain the logic CTL^*_{lp}, a well-known linear-past and equally expressive extension of CTL^* [11, 12]. We also consider the finitary variant of hybrid CTL^*_{lp}, where the path quantifier $\exists$ is replaced with the finitary path quantifier $\exists_f$. This logic corresponds to an extension of finitary CTL^* and its semantics is similar to that of hybrid CTL^*_{lp} with the exception that path quantification ranges over the *finite* paths (tracks) that start at the root and visit the current node of the computation tree.

In the following we will use the fragment of hybrid CTL^*_{lp} consisting of *well-formed* formulas, namely, formulas φ where: *(1) each subformula $\exists\psi$ of φ has at most one free variable; (2) each subformula $\exists\psi(x)$ of φ having x as free variable occurs in φ in the context* $(\mathsf{F}^- x) \wedge \exists\psi(x)$. Intuitively, for each state subformula $\exists\psi$, the unique free variable (if any) refers to ancestors of the current node in the computation tree. The notion of well-formed formula of finitary hybrid CTL^*_{lp} is similar: the path quantifier $\exists$ is replaced by its finitary version $\exists_f$. The well-formedness constraint ensures that a formula captures only branching regular requirements. As an example, the formula $\exists\mathsf{F}{\downarrow}x.\mathsf{G}^-(\neg\mathsf{X}^-\top \rightarrow \forall\mathsf{F}(x \wedge p))$ is *not* well-formed and requires that there is a level of the computation tree such that each node in the level satisfies p. This represents a well-known non-regular context-free branching requirement (see, e.g., [2]). We first show that $\mathsf{HS_{lp}}$ can be translated into the well-formed fragment of hybrid CTL^*_{lp} (resp., well-formed fragment of finitary hybrid CTL^*_{lp}). Then we show that this fragment is subsumed by CTL^* (resp., finitary CTL^*).

▶ Proposition 2. Given a $\mathsf{HS_{lp}}$ formula φ, one can construct in linear-time an equivalent well-formed sentence of hybrid CTL^*_{lp} (resp., finitary hybrid CTL^*_{lp}).

Proof. We focus on the translation from $\mathsf{HS_{lp}}$ into the well-formed fragment of hybrid CTL^*_{lp}. The translation from $\mathsf{HS_{lp}}$ into the well-formed fragment of finitary hybrid CTL^*_{lp} is similar. Let φ be a $\mathsf{HS_{lp}}$ formula. The desired hybrid CTL^*_{lp} sentence is given by ${\downarrow}x.\mathsf{G}\, f(\varphi, x)$, where the mapping $f(\varphi, x)$ is homomorphic with respect to the Boolean connectives, and is defined for the atomic propositions and the other modalities as follows (y is a fresh variable):

$$\begin{array}{ll}
f(p, x) & = \mathsf{G}^-((\mathsf{F}^- x) \rightarrow p), \\
f(\langle \mathrm{B} \rangle\, \psi, x) & = \mathsf{X}^-\mathsf{F}^-(f(\psi, x) \wedge \mathsf{F}^- x), \\
f(\langle \overline{\mathrm{B}} \rangle\, \psi, x) & = (\mathsf{F}^- x) \wedge \exists(\mathsf{XF} f(\psi, x)), \\
f(\langle \mathrm{E} \rangle\, \psi, x) & = {\downarrow} y.\mathsf{F}^-\big(x \wedge \mathsf{XF}{\downarrow}x.\mathsf{F}(y \wedge f(\psi, x))\big), \\
f(\langle \overline{\mathrm{E}} \rangle\, \psi, x) & = {\downarrow} y.\mathsf{F}^-\big((\mathsf{XF}x) \wedge {\downarrow}x.\mathsf{F}(y \wedge f(\psi, x))\big).
\end{array}$$

Clearly $\downarrow x.\mathsf{G}\, f(\varphi, x)$ is well-formed. Moreover, let $\mathcal{K}$ be a Kripke structure, $[h, i]$ be an interval of positions, g be a valuation assigning to the variable x the position h, and π be an initial infinite path. By a straightforward induction on the structure of φ, one can show that $\mathcal{K}, \pi, g, i \models f(\varphi, x)$ if and only if $\mathcal{C}(\mathcal{K}), \mathcal{C}(\pi, h, i) \models_{\mathsf{st}} \varphi$, where $\mathcal{C}(\pi, h, i)$ denotes the track of the computation tree $\mathcal{C}(\mathcal{K})$ starting from $\pi[0, h]$ and leading to $\pi[0, i]$. Hence $\mathcal{K}$ is a model of $\downarrow x.\mathsf{G}\, f(\varphi, x)$ if for each initial track ρ of $\mathcal{C}(\mathcal{K})$ we have $\mathcal{C}(\mathcal{K}), \rho \models_{\mathsf{st}} \varphi$. ◀

Let LTL_p be the past extension of LTL, obtained by adding the past modalities X^- and U^-. By exploiting the well-formedness requirement and the well-known separation theorem for LTL_p over finite and infinite words [8] (i.e., any LTL_p formula can be effectively converted into an equivalent Boolean combination of LTL formulas and pure past LTL_p formulas), and proceeding by induction on the nesting depth of path quantifiers, we establish the following result (the proof is given in Appendix C).

▶ Proposition 3. The set of well-formed sentences of hybrid CTL^*_{lp} (resp., finitary hybrid CTL^*_{lp}) has the same expressiveness as CTL^* (resp., finitary CTL^*).

By Corollary 18, and Propositions 2 and 3, we obtain the main result of Section 3.2.

▶ **Theorem 19.** *$\mathsf{CTL}^* \geq \mathsf{HS}_{\mathsf{lp}}$. Moreover, $\mathsf{HS}_{\mathsf{lp}}$ is as expressive as finitary CTL^*.*

3.3 Expressiveness comparison of $\mathsf{HS}_{\mathsf{lin}}$, $\mathsf{HS}_{\mathsf{st}}$ and $\mathsf{HS}_{\mathsf{lp}}$

We first show that $\mathsf{HS}_{\mathsf{st}}$ is *not* subsumed by $\mathsf{HS}_{\mathsf{lp}}$. As a matter of fact we show that $\mathsf{HS}_{\mathsf{st}}$ is sensitive to unwinding, allowing us to discriminate finite Kripke structures having the same computation tree (whereas they are indistinguishable by $\mathsf{HS}_{\mathsf{lp}}$). In particular, let us consider the two finite Kripke structures $\mathcal{K}_1$ and $\mathcal{K}_2$ of Figure 3. Since $\mathcal{K}_1$ and $\mathcal{K}_2$ have the same computation tree, no HS formula φ under the computation-tree-based semantics can distinguish $\mathcal{K}_1$ and $\mathcal{K}_2$, i.e., $\mathcal{K}_1 \models_{\mathsf{lp}} \varphi$ iff $\mathcal{K}_2 \models_{\mathsf{lp}} \varphi$. On the other hand, the requirement "*each state reachable from the initial one where p holds has a predecessor where p holds as well*" can be expressed, under the state-based semantics, by the HS formula $\psi := \langle E \rangle (p \wedge length_1) \rightarrow \langle E \rangle (length_1 \wedge \langle \overline{A} \rangle (p \wedge \neg length_1))$. Clearly $\mathcal{K}_1 \models_{\mathsf{st}} \psi$ but $\mathcal{K}_2 \not\models_{\mathsf{st}} \psi$. Hence we obtain the following result.

▶ Proposition 4. $\mathsf{HS}_{\mathsf{lp}} \not\geq \mathsf{HS}_{\mathsf{st}}$.

Since $\mathsf{HS}_{\mathsf{lp}}$ and finitary CTL^* have the same expressiveness (Theorem 19) and finitary CTL^* is subsumed by $\mathsf{HS}_{\mathsf{st}}$ (Corollary 18), by Proposition 4 we obtain the following result.

▶ **Corollary 20.** *$\mathsf{HS}_{\mathsf{st}}$ is more expressive than $\mathsf{HS}_{\mathsf{lp}}$.*

Let us now consider the CTL formula $\forall\mathsf{G}\exists\mathsf{F}p$ asserting that from each state reachable from the initial one, it is possible to reach a state where p holds. It is well-known that this formula is not LTL-expressible. Thus, by Corollary 13, there is no equivalent HS formula under the trace-based semantics. On the other hand, the requirement $\forall\mathsf{G}\exists\mathsf{F}p$ can be expressed under the state-based (resp., computation-tree-based) semantics by the HS formula $\langle \overline{B} \rangle \langle E \rangle p$.

▶ Proposition 5. $\mathsf{HS}_{\mathsf{lin}} \not\geq \mathsf{HS}_{\mathsf{st}}$ and $\mathsf{HS}_{\mathsf{lin}} \not\geq \mathsf{HS}_{\mathsf{lp}}$.

Next we show that $\mathsf{HS}_{\mathsf{lin}} \not\leq \mathsf{HS}_{\mathsf{st}}$ and $\mathsf{HS}_{\mathsf{lin}} \not\leq \mathsf{HS}_{\mathsf{lp}}$. To this end we establish the following.

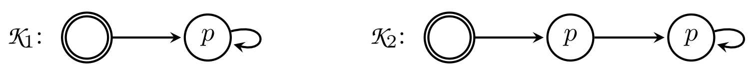


Figure 3 The Kripke structures $\mathcal{K}_1$ and $\mathcal{K}_2$.

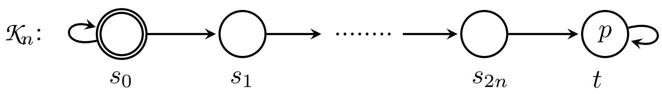


Figure 4 The Kripke structure $\mathcal{K}_n$ with $n \geq 1$.

▶ Proposition 6. The LTL formula $\mathsf{F}\,p$ (equivalent to the CTL formula $\forall\mathsf{F}\,p$) cannot be expressed in either $\mathsf{HS}_{\mathsf{lp}}$ or $\mathsf{HS}_{\mathsf{st}}$.

We prove Proposition 6 for the state-based semantics (for the computation-tree-based semantics the proof is similar). We exhibit two families of Kripke structures $(\mathcal{K}_n)_{n\geq 1}$ and $(\mathcal{M}_n)_{n\geq 1}$ over $\{p\}$ such that for all $n \geq 1$ the LTL formula $\mathsf{F}\,p$ distinguishes $\mathcal{K}_n$ and $\mathcal{M}_n$, and for every HS formula ψ of size at most n, ψ does *not* distinguish $\mathcal{K}_n$ and $\mathcal{M}_n$ under the state-based semantics. Hence the result follows. Fix $n \geq 1$. The Kripke structure $\mathcal{K}_n$ is given in Figure 4. The Kripke structure $\mathcal{M}_n$ is obtained from $\mathcal{K}_n$ by setting as its initial state s_1 instead of s_0. Formally, $\mathcal{K}_n = (\{p\}, S_n, \delta_n, \mu_n, s_0)$ and $\mathcal{M}_n = (\{p\}, S_n, \delta_n, \mu_n, s_1)$, where $S_n = \{s_0, s_1, \ldots, s_{2n}, t\}$, $\delta_n = \{(s_0, s_0), (s_0, s_1), \ldots, (s_{2n-1}, s_{2n}), (s_{2n}, t), (t, t)\}$, $\mu(s_i) = \emptyset$ for all $0 \leq i \leq 2n$, and $\mu(t) = \{p\}$. Clearly $\mathcal{K}_n \not\models \mathsf{F}p$ and $\mathcal{M}_n \models \mathsf{F}p$.

We say that a HS formula ψ is *balanced* if, for each subformula $\langle B\rangle\,\theta$ (resp., $\langle \overline{B}\rangle\,\theta$), θ is of the form $\theta_1 \wedge \theta_2$ with $|\theta_1| = |\theta_2|$. By using conjunctions of $\top$, one can trivially convert a HS formula ψ into a balanced HS formula which is equivalent to ψ under any of the considered HS semantic variants. Lemma 21 is proved in Appendix D: by such a lemma and the fact that, for each $n \geq 1$, $\mathcal{K}_n \not\models \mathsf{F}p$ and $\mathcal{M}_n \models \mathsf{F}p$, we get a proof of Proposition 6.

▶ **Lemma 21.** *For all $n \geq 1$ and balanced HS formulas ψ s.t. $|\psi| \leq n$, $\mathcal{K}_n \models_{\mathsf{st}} \psi$ iff $\mathcal{M}_n \models_{\mathsf{st}} \psi$.*

By Propositions 5–6, we obtain the following result.

▶ **Corollary 22.** *$\mathsf{HS}_{\mathsf{lin}}$ and $\mathsf{HS}_{\mathsf{st}}$ (resp., $\mathsf{HS}_{\mathsf{lp}}$) are expressively incomparable.*

The proved results also allow us to establish the expressiveness relations between $\mathsf{HS}_{\mathsf{st}}$, $\mathsf{HS}_{\mathsf{lp}}$ and the standard branching temporal logics CTL and CTL*.

▶ **Corollary 23.** *(1) $\mathsf{HS}_{\mathsf{st}}$ and CTL* (resp., CTL) are expressively incomparable; (2) $\mathsf{HS}_{\mathsf{lp}}$ and finitary CTL* are less expressive than CTL*; (3) $\mathsf{HS}_{\mathsf{lp}}$ and CTL are expressively incomparable.*

Proof. (Point 1) By Proposition 6 and the fact that CTL* is not sensitive to unwinding.
(Point 2) By Theorem 19, $\mathsf{HS}_{\mathsf{lp}}$ is subsumed by CTL*, and $\mathsf{HS}_{\mathsf{lp}}$ and finitary CTL* have the same expressiveness. Hence, by Proposition 6, the result follows.
(Point 3) By Proposition 6, it suffices to show that there exists a $\mathsf{HS}_{\mathsf{lp}}$ formula which cannot be expressed in CTL. Let us consider the CTL* formula $\varphi := \exists\big(((p_1 \mathsf{U} p_2) \vee (q_1 \mathsf{U} q_2))\,\mathsf{U}\,r\big)$ over the set of propositions $\{p_1, p_2, q_1, q_2, r\}$. It is shown in [7] that φ cannot be expressed in CTL. Clearly if we replace the path quantifier $\exists$ in φ with the finitary path quantifier $\exists_f$, we obtain an equivalent formula of finitary CTL*. Thus, since $\mathsf{HS}_{\mathsf{lp}}$ and finitary CTL* have the same expressiveness (Theorem 19), the result follows. ◀

4 Conclusions and future work

In this paper, we have studied three semantic variants, namely, $\mathsf{HS}_{\mathsf{st}}$, $\mathsf{HS}_{\mathsf{lp}}$, and $\mathsf{HS}_{\mathsf{lin}}$, of the interval temporal logic HS, comparing their expressiveness to that of the standard temporal logics LTL, CTL, finitary CTL*, and CTL*. The reported results imply the decidability of the model checking problem for $\mathsf{HS}_{\mathsf{lp}}$ and $\mathsf{HS}_{\mathsf{lin}}$; the related complexity issues will be studied in the future work. Moreover, we shall investigate how the expressiveness changes when the homogeneity assumption is relaxed.

References

1. J. F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843, 1983.
2. R. Alur, P. Cerný, and S. Zdancewic. Preserving secrecy under refinement. In *ICALP*, LNCS 4052, pages 107–118. Springer, 2006.
3. P. Blackburn and J. Seligman. What are hybrid languages? In *AiML*, pages 41–62. CSLI Publications, 1998.
4. L. Bozzelli, A. Molinari, A. Montanari, A. Peron, and P. Sala. Interval Temporal Logic Model Checking: the Border Between Good and Bad HS Fragments. In *IJCAR*, LNAI 9706, pages 389–405. Springer, 2016. `doi:10.1007/978-3-319-40229-1_27`.
5. L. Bozzelli, A. Molinari, A. Montanari, A. Peron, and P. Sala. Model Checking the Logic of Allen's Relations Meets and Started-by is $\mathbf{P}^{\mathbf{NP}}$-Complete. In *GandALF*, pages 76–90, 2016. `doi:10.4204/EPTCS.226.6`.
6. D. Bresolin, D. Della Monica, V. Goranko, A. Montanari, and G. Sciavicco. The dark side of interval temporal logic: marking the undecidability border. *Annals of Mathematics and Artificial Intelligence*, 71(1-3):41–83, 2014.
7. E. A. Emerson and J. Y. Halpern. "Sometimes" and "not never" revisited: on branching versus linear time temporal logic. *Journal of the ACM*, 33(1):151–178, 1986.
8. D. M. Gabbay. The declarative past and imperative future: Executable temporal logic for interactive systems. In *Temporal Logic in Specification*, LNCS 398, pages 409–448. Springer, 1987.
9. J. Y. Halpern and Y. Shoham. A propositional modal logic of time intervals. *Journal of the ACM*, 38(4):935–962, 1991.
10. H. Kamp. *Tense Logic and the Theory of Linear Order*. PhD thesis, Ucla, 1968.
11. O. Kupferman and A. Pnueli. Once and for all. In *LICS*, pages 25–35. IEEE Computer Society, 1995.
12. O. Kupferman, A. Pnueli, and M. Y. Vardi. Once and for all. *J. Comput. Syst. Sci.*, 78(3):981–996, 2012.
13. K. Lodaya. Sharpening the undecidability of interval temporal logic. In *ASIAN*, LNCS 1961, pages 290–298. Springer, 2000.
14. A. Lomuscio and J. Michaliszyn. An epistemic Halpern-Shoham logic. In *IJCAI*, pages 1010–1016, 2013.
15. A. Lomuscio and J. Michaliszyn. Decidability of model checking multi-agent systems against a class of EHS specifications. In *ECAI*, pages 543–548, 2014.
16. A. Lomuscio and J. Michaliszyn. Model checking multi-agent systems against epistemic HS specifications with regular expressions. In *KR*, pages 298–308. AAAI Press, 2016.
17. J. Marcinkowski and J. Michaliszyn. The undecidability of the logic of subintervals. *Fundamenta Informaticae*, 131(2):217–240, 2014.
18. A. Molinari, A. Montanari, A. Murano, G. Perelli, and A. Peron. Checking interval properties of computations. *Acta Informatica*, 2016. Accepted for publication. `doi:10.1007/s00236-015-0250-1`.
19. A. Molinari, A. Montanari, and A. Peron. Complexity of ITL model checking: some well-behaved fragments of the interval logic HS. In *TIME*, pages 90–100, 2015. `doi:10.1109/TIME.2015.12`.
20. A. Molinari, A. Montanari, and A. Peron. A model checking procedure for interval temporal logics based on track representatives. In *CSL*, pages 193–210, 2015. `doi:10.4230/LIPIcs.CSL.2015.193`.
21. A. Molinari, A. Montanari, A. Peron, and P. Sala. Model Checking Well-Behaved Fragments of HS: the (Almost) Final Picture. In *KR*, pages 473–483, 2016.

22 A. Montanari, A. Murano, G. Perelli, and A Peron. Checking interval properties of computations. In *TIME*, pages 59–68, 2014.

23 B. Moszkowski. *Reasoning About Digital Circuits.* PhD thesis, Stanford University, CA, 1983.

24 A. Pnueli. The temporal logic of programs. In *FOCS*, pages 46–57. IEEE, 1977.

25 I. Pratt-Hartmann. Temporal prepositions and their logic. *Artificial Intelligence*, 166(1-2):1–36, 2005.

26 P. Roeper. Intervals and tenses. *Journal of Philosophical Logic*, 9:451–469, 1980.

27 M. Y. Vardi. An automata-theoretic approach to linear temporal logic. In *Logics for concurrency*, pages 238–266. Springer, 1996.

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

29 T. Wilke. Classifying discrete temporal properties. In *STACS*, LNCS 1563, pages 32–46. Springer, 1999.

Appendix

A Proof of Theorem 11

Recall that we interpret FO formulas φ over infinite paths π of Kripke structures $\mathcal{K} = (\mathcal{AP}, S, \delta, \mu, s_0)$. Given a variable valuation g, the satisfaction relation $(\pi, g) \models \varphi$ is inductively defined as follows (we omit the standard rules for the Boolean connectives):

$$\begin{aligned}
&(\pi, g) \models p \in x && \Leftrightarrow p \in \mu(\pi(g(x))),\\
&(\pi, g) \models x \; op \; y && \Leftrightarrow g(x) \; op \; g(y), \text{ for } op \in \{<, \leq\},\\
&(\pi, g) \models \exists x.\varphi && \Leftrightarrow (\pi, g[x \leftarrow i]) \models \varphi \text{ for some } i \geq 0,
\end{aligned}$$

where $g[x \leftarrow i](x) = i$ and $g[x \leftarrow i](y) = g(y)$ for $y \neq x$. Notice that the satisfaction relation depends only on the values assigned to the variables occurring free in the given formula φ. Now we prove Theorem 11.

▶ **Theorem 11.** *LTL* $\geq$ *HS*$_{\mathsf{lin}}$.

Proof. We first inductively define a mapping h assigning to each triple (φ, x, y), consisting of a HS formula φ and two distinct positions variables x, y, a FO formula having as free variables x and y. The mapping h is homomorphic with respect to the Boolean connectives, and is defined for atomic propositions and modal operators as follows (z is a fresh position variable):

$$\begin{aligned}
&h(p, x, y) && = \forall z.((z \geq x \wedge z \leq y) \rightarrow p \in z),\\
&h(\langle E \rangle \psi, x, y) && = \exists z.(z > x \wedge z \leq y \wedge h(\psi, z, y)),\\
&h(\langle B \rangle \psi, x, y) && = \exists z.(z \geq x \wedge z < y \wedge h(\psi, x, z)),\\
&h(\langle \overline{E} \rangle \psi, x, y) && = \exists z.(z < x \wedge h(\psi, z, y)),\\
&h(\langle \overline{B} \rangle \psi, x, y) && = \exists z.(z > y \wedge h(\psi, x, z)).
\end{aligned}$$

Given a Kripke structure $\mathcal{K}$, an infinite path π, an interval of positions $[i, j]$, and an HS$_{\mathsf{lin}}$ formula ψ, by a straightforward induction on the structure of ψ, we can show that $\mathcal{IS}_{\mathcal{K},\pi}, [i, j] \models \psi$ iff $(\pi, g) \models h(\psi, x, y)$ for any valuation such that $g(x) = i$ and $g(y) = j$. Now, let us consider the FO sentence $h(\psi)$ given by $\exists x((\forall z. z \geq x) \wedge \forall y. h(\psi, x, y))$. Clearly $\mathcal{K} \models_{\mathsf{lin}} \psi$ iff for each initial infinite path π of $\mathcal{K}$, it holds that $\pi \models h(\psi)$. By Proposition 1, it follows that one can construct an LTL formula $h'(\psi)$ such that $h'(\psi)$ in LTL is equivalent to ψ in HS$_{\mathsf{lin}}$. ◀

B Proof of Theorem 16

By Theorem 15, we just need to prove the following.

▶ **Theorem 24.** *The class of languages of finite words definable by* BE *formulas is* LTL*-closed.*

Since the class of BE-definable languages is obviously closed under language complementation and language intersection, Theorem 24 directly follows from Definition 14 and the following two Lemmata 25 and 26.

▶ **Lemma 25.** *Let Σ be a finite alphabet, $b \in \Sigma$, $\Gamma = \Sigma \setminus \{b\}$, $L \subseteq \Gamma^+$, and ψ be a* BE *formula over Γ such that $L_{act}(\psi) = L$. Then, one can construct* BE *formulas capturing under the action-based semantics the languages bL, $\Sigma^* bL$, $\Sigma^* b(L + \varepsilon)$, Lb, $Lb\Sigma^*$, $(L + \varepsilon)b\Sigma^*$, and bLb.*

Proof. We focus on the languages bL, $\Sigma^* bL$, and bLb (for the other languages the proof is similar). Let ψ be a BE formula over Γ such that $L_{act}(\psi) = L$.

Proof for the language bL: we construct by structural induction on ψ a BE formula $h_b(\psi)$ as follows. The mapping h_b is homomorphic with respect to the Boolean connectives, and is defined for the atomic actions in Γ and modalities $\langle E \rangle$ and $\langle B \rangle$ as follows:

- for all $a \in \Gamma$, $h_b(a) = a \vee (\langle B \rangle\, b \wedge \langle E \rangle\, a \wedge [E]a)$;
- $h_b(\langle B \rangle\, \theta) = (\langle B \rangle\, h_b(\theta) \wedge \neg \langle B \rangle\, b) \vee \langle B \rangle (h_b(\theta) \wedge \langle B \rangle\, b)$;
- $h_b(\langle E \rangle\, \theta) = (\langle E \rangle\, h_b(\theta) \wedge \neg \langle B \rangle\, b) \vee (\langle B \rangle\, b \wedge \langle E \rangle \langle E \rangle\, h_b(\theta))$.

We can show by a straightforward structural induction on ψ that the following fact holds.

Claim 1: Let $u \in \Gamma^+$, $u' = bu$, and $|u| = n+1$. Then, for all $i, j \in [0, n]$ with $i \leq j$, $u[i,j] \in L_{act}(\psi)$ iff $u'[\hat{\imath}, j+1] \in L_{act}(h_b(\psi))$ where $\hat{\imath} = i$ if $i = 0$, and $\hat{\imath} = i+1$ otherwise.

By Claim 1, for each $u \in \Gamma^+$, $u \in L_{act}(\psi)$ iff $bu \in L_{act}(h_b(\psi))$. Therefore the formula $(\neg length_1 \wedge \langle B \rangle\, b \wedge [E](\neg b \wedge [B]\neg b)) \wedge h_b(\psi)$ captures the language $bL_{act}(\psi)$.

Proof for the language $\Sigma^* bL$: by the proof given for the language bL, where $L \subseteq \Gamma^+$, one can construct a BE formula φ capturing bL. Hence a BE formula capturing $\Sigma^* bL$ is $\varphi \vee \langle E \rangle\, \varphi$.

Proof for the language bLb: by the proof given for the language bL, where $L \subseteq \Gamma^+$, one can build a BE formula φ capturing bL. We construct by structural induction on φ a BE formula $k_b(\varphi)$ as follows. The mapping k_b is homomorphic with respect to the Boolean connectives, and is defined for the atomic actions in Σ and modalities $\langle E \rangle$ and $\langle B \rangle$ as follows:

- for all $a \in \Gamma$, $k_b(a) = a \vee (\langle E \rangle\, b \wedge \langle B \rangle\, a \wedge [B]a)$;
- $k_b(b) = b$;
- $k_b(\langle B \rangle\, \theta) = (\langle B \rangle\, k_b(\theta) \wedge \neg \langle E \rangle\, b) \vee (\langle E \rangle\, b \wedge \langle B \rangle \langle B \rangle\, k_b(\theta))$.
- $k_b(\langle E \rangle\, \theta) = (\langle E \rangle\, k_b(\theta) \wedge \neg \langle E \rangle\, b) \vee \langle E \rangle (k_b(\theta) \wedge \langle E \rangle\, b)$.

By a straightforward structural induction on φ, we can show that the following fact holds.

Claim 2: Let $u \in \Gamma^+$ and $|bu| = n+1$. Then, for all $i, j \in [0, n]$ with $i \leq j$, $bu[i,j] \in L_{act}(\varphi)$ iff $bub[i, \hat{\jmath}] \in L_{act}(k_b(\varphi))$ where $\hat{\jmath} = j$ if $j < n$, and $\hat{\jmath} = n+1$ otherwise.

By Claim 2, for each $u \in \Gamma^+$, $bu \in L_{act}(\varphi)$ iff $bub \in L_{act}(k_b(\varphi))$. Thus the formula $(\neg length_1 \wedge \neg length_2 \wedge \langle B \rangle\, b \wedge \langle E \rangle\, b \wedge [E][B]\neg b) \wedge h_b(\psi)$ captures the language $L_{act}(\varphi)b$, and this concludes the proof of the lemma. ◀

▶ **Lemma 26.** *Let Σ and Δ be finite alphabets, $b \in \Sigma$, $\Gamma = \Sigma \setminus \{b\}$, $U_0 = \Gamma^* b$, $h_0 : U_0 \to \Delta$ and $h : U_0^+ \to \Delta^+$ be defined by $h(u_0 u_1 \dots u_n) = h_0(u_0) \dots h_0(u_n)$. Assume that, for each $d \in \Delta$, there is a* BE *formula capturing the language $L_d = \{u \in \Gamma^+ \mid h_0(ub) = d\}$. Then for each* BE *formula φ over Δ, one can construct a* BE *formula over Σ capturing the language $\Gamma^* b h^{-1}(L_{act}(\varphi))\Gamma^*$.*

Proof. By hypothesis and Lemma 25, for each $d \in \Delta$ there exists a BE formula θ_d over Σ capturing the language $bL_d b$, where $L_d = \{u \in \Gamma^+ \mid h_0(ub) = d\}$. Hence evidently there is a BE formula $\hat{\theta}_d$ over Σ capturing the language $b\hat{L}_d b$, where $\hat{L}_d = \{u \in \Gamma^* \mid h_0(ub) = d\}$ (note that $L_d = \hat{L}_d \setminus \{\varepsilon\}$). Let φ be a BE formula over Δ. By structural induction over φ, we construct a BE formula φ^+ over Σ such that $L_{act}(\varphi^+) = \Gamma^* b h^{-1}(L_{act}(\varphi))\Gamma^*$. The formula φ^+ is defined as follows:

- $\varphi = d$ with $d \in \Delta$. We have that $L_{act}(d) = d^+$ and $\Gamma^* b h^{-1}(L_{act}(d))\Gamma^*$ is the set of finite words in $\Gamma^* b \Sigma^* b \Gamma^*$ such that each subword $u[i,j]$ of u which is in $b\Gamma^* b$ is in $b\hat{L}_d b$ as well.

Thus φ^+ is defined as follows, where $\psi_b := \neg length_1 \wedge \langle \mathrm{B} \rangle\, b \wedge \langle \mathrm{E} \rangle\, b \wedge [E][B]\neg b$ captures the words in $b\Gamma^* b$:

$$\varphi^+ = \langle \mathrm{G} \rangle(\psi_b \wedge \hat{\theta}_d) \wedge [G](\psi_b \rightarrow \hat{\theta}_d).$$

- $\varphi = \neg\theta$. We have that

$$\Gamma^* b h^{-1}(L_{act}(\varphi))\Gamma^* = \Gamma^* b \Sigma^* b \Gamma^* \cap \overline{\Gamma^* b h^{-1}(L_{act}(\theta))\Gamma^*}.$$

  Thus φ^+ is given by $\neg\theta^+ \wedge \langle \mathrm{G} \rangle\, \psi_b$, where ψ_b has been defined in the previous case.
- $\varphi = \theta \wedge \psi$. We take $\varphi^+ = \theta^+ \wedge \psi^+$, and the correctness of the construction easily follows.
- $\varphi = \langle \mathrm{B} \rangle\, \theta$. Clearly $\Gamma^* b h^{-1}(L_{act}(\langle \mathrm{B} \rangle\, \theta))\Gamma^*$ is the set of finite words over Σ featuring a proper prefix in $\Gamma^* b h^{-1}(L_{act}(\theta))\Sigma^* b$. Thus φ^+ is given by:

$$\varphi^+ = \langle \mathrm{B} \rangle(\langle \mathrm{E} \rangle\, b \wedge \langle \mathrm{B} \rangle(\theta^+ \wedge \langle \mathrm{E} \rangle\, b)).$$

- $\varphi = \langle \mathrm{E} \rangle\, \theta$. Clearly $\Gamma^* b h^{-1}(L_{act}(\langle \mathrm{E} \rangle\, \theta))\Gamma^*$ is the set of finite words over Σ featuring a proper suffix in $b\Sigma^* b h^{-1}(L_{act}(\theta))\Gamma^*$. Thus φ^+ is given by:

$$\varphi^+ = \langle \mathrm{E} \rangle(\langle \mathrm{B} \rangle\, b \wedge \langle \mathrm{E} \rangle(\theta^+ \wedge \langle \mathrm{B} \rangle\, b)).$$

The proof of the lemma is complete. ◀

C Proof of Proposition 3

In this section, we show that the well-formed fragment of hybrid CTL^*_{lp} (resp., finitary hybrid CTL^*_{lp}) is not more expressive than CTL^* (resp., finitary CTL^*). Here we focus on the well-formed fragment of hybrid CTL^*_{lp} (the proof for the finitary variant is similar).

We need additional definitions and preliminary results. A *pure past* LTL_p formula is a LTL_p formula which does not contain occurrences of future temporal modalities. Given two formulas φ and φ' of hybrid CTL^*_{lp}, φ and φ' are *congruent* if for every Kripke structure $\mathcal{K}$, initial infinite path π, valuation g, and current position i, $\mathcal{K}, \pi, g, i \models \varphi$ iff $\mathcal{K}, \pi, g, i \models \varphi'$. As usual, for a formula φ of hybrid CTL^*_{lp} with one free variable x, we also write $\varphi(x)$; moreover, since the satisfaction relation depends only on the variables occurring free in the given formula, for $\varphi(x)$ we use the notation $\mathcal{K}, \pi, i \models \varphi(x \leftarrow h)$ to mean that $\mathcal{K}, \pi, g, i \models \varphi$ for any valuation g assigning h to the unique free variable x. For a formula φ of hybrid CTL^*_{lp}, let $\exists SubF(\varphi)$ be the set of subformulas of φ of the form $\exists\psi$ which do not occur in the scope of the path quantifier $\exists$. We now introduce the notion of *simple* hybrid CTL^*_{lp} formula.

▶ **Definition 27.** Given a variable x, a *simple* hybrid CTL^*_{lp} formula ψ with respect to x is a hybrid CTL^*_{lp} formula satisfying the following syntactical requirements:
- x is the unique variable occurring in ψ;
- ψ does *not* contain occurrences of the binder modalities and past temporal modalities;
- $\exists SubF(\psi)$ consists of CTL^* formulas.

Intuitively, a *simple* hybrid CTL^*_{lp} formula ψ with respect to x over $\mathcal{AP}$ can be seen as a CTL^* formula over the set of propositions $\mathcal{AP} \cup \{x\}$. We make now the following observation.

▶ **Lemma 28.** *Let ψ be a* simple *hybrid CTL^*_{lp} formula with respect to x. Then $(\mathsf{F}^- x) \wedge \psi$ is congruent to $(\mathsf{F}^- x) \wedge \xi$, where ξ is a Boolean combination of the atomic formula x and CTL^* formulas.*

Proof. Let ψ be a *simple* hybrid CTL^*_{lp} formula with respect to x. We denote by $\widehat{\psi}$ the CTL^* formula obtained from ψ by replacing each occurrence of x in ψ with *false*. The proof is by structural induction on ψ. The base case is obvious. As for the inductive step, ψ is a Boolean combination of simple hybrid CTL^*_{lp} formulas θ, where θ is either an atomic proposition, or the variable x, or a CTL^* formula, or a simple formula of the form $\mathsf{X}\theta_1$ or $\theta_1\mathsf{U}\theta_2$. Thus we just need to consider the cases where $\theta = \mathsf{X}\theta_1$ or $\theta_1\mathsf{U}\theta_2$. For $\theta = \mathsf{X}\theta_1$, the result follows from the fact that $(\mathsf{F}^- x) \wedge \theta$ is congruent to $(\mathsf{F}^- x) \wedge \mathsf{X}\widehat{\theta_1}$. For $\theta = \theta_1\mathsf{U}\theta_2$, the result follows from the inductive hypothesis and the fact that $(\mathsf{F}^- x) \wedge \theta$ is congruent to $(\mathsf{F}^- x) \wedge (\theta_2 \vee \mathsf{X}(\widehat{\theta_1\mathsf{U}\theta_2}))$. ◀

Then we deduce the following crucial technical result by exploiting the well-known separation theorem for LTL_p over infinite words [8].

▶ **Lemma 29.** *Let $(\mathsf{F}^- x) \wedge \exists\varphi(x)$ (resp., $\exists\varphi$) be a well-formed formula (resp., well-formed sentence) of hybrid CTL^*_{lp} such that $\exists SubF(\varphi)$ consists of CTL^* formulas. Then, $(\mathsf{F}^- x) \wedge \exists\varphi(x)$ (resp., $\exists\varphi$) is congruent to a well-formed formula of hybrid CTL^*_{lp} which is a Boolean combination of CTL^* formulas and formulas that correspond to* pure past *LTL_p formulas over the set of atomic propositions given by $\mathcal{AP} \cup \exists SubF(\varphi) \cup \{x\}$ (resp., $\mathcal{AP} \cup \exists SubF(\varphi)$).*

Proof. We focus on well-formed formulas of the form $(\mathsf{F}^- x) \wedge \exists\varphi(x)$ (the case of well-formed sentences of the form $\exists\varphi$ is similar). Let $\overline{\mathcal{AP}} = \mathcal{AP} \cup \exists SubF(\varphi) \cup \{x\}$. By hypothesis $\exists SubF(\varphi)$ consists of CTL^* formulas. Given a Kripke structure $\mathcal{K} = (\mathcal{AP}, S, \delta, \mu, s_0)$, an initial infinite path π, and $h \geq 0$, we denote by $\pi_{\overline{\mathcal{AP}},h}$ the infinite word over $2^{\overline{\mathcal{AP}}}$ defined as follows for every position $i \geq 0$:

- $\pi_{\overline{\mathcal{AP}},h}(i) \cap \mathcal{AP} = \mu(\pi(i))$;
- $\pi_{\overline{\mathcal{AP}},h}(i) \cap \exists SubF(\varphi) = \{\psi \in \exists SubF(\varphi) \mid \mathcal{K}, \pi, i \models \psi\}$;
- $x \in \pi_{\overline{\mathcal{AP}},h}(i)$ iff $i = h$.

By using a fresh position variable *present*, which intuitively represents the current position, the formula $\varphi(x)$ can be easily converted into a FO formula $\varphi_{\mathsf{FO}}(present)$ over $\overline{\mathcal{AP}}$ having *present* as unique free variable, such that for all Kripke structures $\mathcal{K}$, initial infinite paths π, and positions i and h:

$$\mathcal{K}, \pi, i \models \varphi(x \leftarrow h) \text{ iff } \pi_{\overline{\mathcal{AP}},h} \models \varphi_{\mathsf{FO}}(present \leftarrow i). \tag{1}$$

By the well-known separation theorem for LTL_p [8] and the equivalence of FO and LTL_p over infinite words, starting from the FO formula $\varphi_{\mathsf{FO}}(present)$, one can construct an LTL_p formula φ_{LTL_p} over $\overline{\mathcal{AP}}$ of the form

$$\varphi_{\mathsf{LTL}_p} := \bigvee_{i \in I} (\varphi_{p,i}(x) \wedge \varphi_{f,i}(x)) \tag{2}$$

such that $\varphi_{p,i}(x)$ is a pure past LTL_p formula, $\varphi_{f,i}(x)$ is a LTL formula, and for all infinite words w over $2^{\overline{\mathcal{AP}}}$ and $i \geq 0$,

$$w, i \models \varphi_{\mathsf{LTL}_p} \text{ iff } w \models \varphi_{\mathsf{FO}}(present \leftarrow i). \tag{3}$$

The LTL_p formula φ_{LTL_p} over $\overline{\mathcal{AP}}$ corresponds to a hybrid CTL^*_{lp} formula over $\mathcal{AP}$. By definition of the infinite words $\pi_{\overline{\mathcal{AP}},h}$, one can easily show by structural induction that for all Kripke structures $\mathcal{K}$, initial infinite paths π, and positions i and h:

$$\pi_{\overline{\mathcal{AP}},h}, i \models \varphi_{\mathsf{LTL}_p} \text{ iff } \mathcal{K}, \pi, i \models \varphi_{\mathsf{LTL}_p}(x \leftarrow h). \tag{4}$$

Thus by Equations (1), (3), and (4), we obtain that

- φ and φ_{LTL_p} are congruent.

Since in Equation (2), for each $i \in I$, $\varphi_{p,i}(x)$ is a pure past LTL_p formula over $\overline{\mathcal{AP}}$, $\exists\varphi_{p,i}(x)$ is congruent to $\varphi_{p,i}(x)$. Hence we obtain that:

- $(\mathsf{F}^- x) \wedge \exists\varphi(x)$ is congruent to $(\mathsf{F}^- x) \wedge \bigvee_{i\in I}(\varphi_{p,i}(x) \wedge \exists\varphi_{f,i}(x))$.

Since $\varphi_{f,i}(x)$ is a *simple* hybrid CTL^*_{lp} formula with respect to x, and $\exists x$ (resp., $\exists\neg x$) is congruent to x (resp., $\neg x$), by applying Lemma 28 we obtain that $(\mathsf{F}^- x) \wedge \exists\varphi(x)$ is congruent to a formula of the form $(\mathsf{F}^- x) \wedge \bigvee_{i\in I}(\psi_{p,i}(x) \wedge \exists\psi_i)$, where ψ_i is a CTL^* formula and $\psi_{p,i}(x)$ corresponds to a *pure past* LTL_p formula over the set of atomic propositions given by $\overline{\mathcal{AP}} = \mathcal{AP} \cup \exists \mathit{SubF}(\varphi) \cup \{x\}$. This concludes the proof of the lemma. ◀

By applying Lemma 29, we deduce the following corollary.

▶ **Corollary 30.** *Let $(\mathsf{F}^- x) \wedge \exists\varphi(x)$ (resp., $\exists\varphi$) be a well-formed formula (resp., well-formed sentence) of hybrid CTL^*_{lp}. Then there exists a finite set $\mathcal{H}$ of CTL^* formulas of the form $\exists\psi$, such that $(\mathsf{F}^- x) \wedge \exists\varphi(x)$ (resp., $\exists\varphi$) is congruent to a well-formed formula of hybrid CTL^*_{lp} which is a Boolean combination of CTL^* formulas and formulas that correspond to* pure past *LTL_p formulas over the set of atomic propositions given by $\mathcal{AP} \cup \mathcal{H} \cup \{x\}$ (resp., $\mathcal{AP} \cup \mathcal{H}$).*

Proof. We focus on well-formed formulas of the form $(\mathsf{F}^- x) \wedge \exists\varphi(x)$ (the case of well-formed sentences of the form $\exists\varphi$ is similar). The proof is by induction on the nesting depth of the path quantifier $\exists$ in $\varphi(x)$. In the base case $\exists \mathit{SubF}(\varphi) = \emptyset$, thus we can apply Lemma 29, and the result follows with $\mathcal{H} = \emptyset$. As for the inductive step, let us consider the non-empty set of formulas $\exists \mathit{SubF}(\varphi)$. Let $\exists\psi \in \exists \mathit{SubF}(\varphi)$. Since $(\mathsf{F}^- x) \wedge \exists\varphi(x)$ is well-formed, either ψ is a sentence, or ψ has a unique free variable y and $\exists\psi(y)$ occurs in $\varphi(x)$ in the context $(\mathsf{F}^- y) \wedge \exists\psi(y)$. Assume that the latter case holds (the former is similar). By the inductive hypothesis we can apply the corollary to $(\mathsf{F}^- y) \wedge \exists\psi(y)$. Hence there exists a finite set $\mathcal{H}'$ of CTL^* formulas of the form $\exists\theta$, such that $(\mathsf{F}^- y) \wedge \exists\psi(y)$ is congruent to a well-formed formula of hybrid CTL^*_{lp}, say $\xi(y)$, which is a Boolean combination of CTL^* formulas and formulas that correspond to *pure past* LTL_p formulas over the set of atomic propositions given by $\mathcal{AP} \cup \mathcal{H}' \cup \{y\}$. By replacing each occurrence of $(\mathsf{F}^- y) \wedge \exists\psi(y)$ in $\varphi(x)$ with $\xi(y)$, and repeating the procedure for all the formulas in $\exists \mathit{SubF}(\varphi)$, we obtain a well-formed formula of hybrid CTL^*_{lp} of the form $(\mathsf{F}^- x) \wedge \exists\theta(x)$ which is congruent to $(\mathsf{F}^- x) \wedge \exists\varphi(x)$ (notice that the congruence relation is closed under substitution) and such that $\exists \mathit{SubF}(\theta)$ consists of CTL^* formulas. At this point we can apply Lemma 29 and the result follows. ◀

▶ Proposition 7. Well-formed hybrid CTL^*_{lp} has the same expressiveness as CTL^*.

Proof. Let φ be a well-formed sentence of hybrid CTL^*_{lp}. We construct a CTL^* formula which is equivalent to φ, hence the result follows. Clearly φ is equivalent to $\neg\exists\neg\varphi$. Thus, since $\neg\exists\neg\varphi$ is well-formed, by applying Corollary 30, one can convert $\neg\exists\neg\varphi$ into a congruent hybrid CTL^*_{lp} formula which is a Boolean combination of CTL^* formulas and formulas θ which can be seen as pure past LTL_p formulas over the set of propositions $\mathcal{AP} \cup \mathcal{H}$, where $\mathcal{H}$ is a set of CTL^* formulas of the form $\exists\psi$. Since the past temporal modalities in such LTL_p formulas θ refer to the initial position of the initial infinite paths, one can replace θ with an equivalent CTL^* formula $f(\theta)$, where the mapping f is inductively defined as follows:

- $f(p) = p$ for all $p \in \mathcal{AP} \cup \mathcal{H}$;
- f is homomorphic w.r.t. the Boolean connectives;
- $f(\mathsf{X}^-\theta) = \bot$ and $f(\theta_1 \mathsf{U}^- \theta_2) = f(\theta_2)$.

At the end we have obtained a CTL* formula which is equivalent to $\neg\exists\neg\varphi$, and the proof is complete. ◀

By an easy adaptation of the proof of Proposition 7 in which we now exploit the well-known separation theorem for LTL_p over finite words [8], we establish the following result.

▶ Proposition 8. The set of well-formed sentences of finitary hybrid CTL^*_{lp} has the same expressiveness as finitary CTL*.

D Proof of Lemma 21

▶ **Lemma 21.** *For all $n \geq 1$ and balanced HS formulas ψ s.t. $|\psi| \leq n$, $\mathcal{K}_n \models_{\mathsf{st}} \psi$ iff $\mathcal{M}_n \models_{\mathsf{st}} \psi$.*

Fix $n \geq 1$. In order to prove Lemma 21 for n, we need some additional definitions. Let ρ be a track of $\mathcal{K}_n$ (note that $\mathcal{K}_n$ and $\mathcal{M}_n$ feature the same tracks). By construction ρ is of the form $\rho' \cdot \rho''$, where ρ' is a (possibly empty) track visiting only states where p does not hold, and ρ'' is a (possibly empty) track visiting only the state t, where p holds. We say that ρ' (resp., ρ'') is the $\emptyset$-part (resp., p-part) of ρ. Let $N_\emptyset(\rho)$, $N_p(\rho)$, and $D_p(\rho)$ be the natural numbers defined as follows:

- $N_\emptyset(\rho) = |\rho'|$ (the length of the $\emptyset$-part of ρ);
- $N_p(\rho) = |\rho''|$ (the length of the p-part of ρ);
- $D_p(\rho) = 0$ if $N_p(\rho) > 0$ (i.e., $\mathrm{lst}(\rho) = t$); otherwise $D_p(\rho)$ is the length of the minimal track starting from $\mathrm{lst}(\rho)$ and leading to s_{2n}. Notice that by construction, $D_p(\rho)$ is well defined and $0 \leq D_p(\rho) \leq 2n + 1$.

By construction, the following property holds.

▶ Remark. For all tracks ρ and ρ' of $\mathcal{K}_n$, if $D_p(\rho) = D_p(\rho')$, then $\mathrm{lst}(\rho) = \mathrm{lst}(\rho')$.

Now, for each $h \in [1, n]$, we introduce the notion of *h-compatibility* between tracks of $\mathcal{K}_n$. Intuitively, this notion allows us to capture the properties which make two tracks indistinguishable under the state-based semantics by means of balanced HS formulas of size at most h.

▶ **Definition 31** (h-compatibility)**.** Let $h \in [1, n]$. Two tracks ρ and ρ' of $\mathcal{K}_n$ are *h-compatible* if the following conditions hold:
- $N_p(\rho) = N_p(\rho')$;
- either $N_\emptyset(\rho) = N_\emptyset(\rho')$, or $N_\emptyset(\rho) \geq h$ and $N_\emptyset(\rho') \geq h$;
- either $D_p(\rho) = D_p(\rho')$, or $D_p(\rho) \geq h$ and $D_p(\rho') \geq h$.

We denote by $R(h)$ the binary relation over the set of tracks of $\mathcal{K}_n$ such that $(\rho, \rho') \in R(h)$ iff ρ and ρ' are h-compatible. Notice that $R(h)$ is an equivalence relation and $R(h) \subseteq R(h-1)$ for all $h \in [2, n]$.

By construction, the following lemma evidently holds.

▶ **Lemma 32.** *For every track ρ of $\mathcal{K}_n$ starting from s_0 (resp., s_1), there exists a track ρ' of $\mathcal{K}_n$ starting from s_1 (resp., s_0) such that $(\rho, \rho') \in R(n)$.*

Then we deduce some crucial properties of the equivalence relation $R(h)$.

▶ **Lemma 33.** *Let $h \in [2, n]$ and $(\rho, \rho') \in R(h)$. The following properties hold:*
1. *for each proper prefix σ of ρ, there exists a proper prefix σ' of ρ' such that $(\sigma, \sigma') \in R(\lfloor \frac{h}{2} \rfloor)$;*
2. *for each track of the form $\rho \cdot \sigma$, where σ is not empty, there exists a track of the form $\rho' \cdot \sigma'$ such that σ' is not empty and $(\rho \cdot \sigma, \rho' \cdot \sigma') \in R(\lfloor \frac{h}{2} \rfloor)$;*

3. *for each proper suffix σ of ρ, there exists a proper suffix σ' of ρ' such that $(\sigma, \sigma') \in R(h-1)$;*
4. *for each track of the form $\sigma \cdot \rho$, where σ is not empty, there exists a track of the form $\sigma' \cdot \rho'$ such that σ' is not empty and $(\sigma \cdot \rho, \sigma' \cdot \rho') \in R(h)$.*

Proof. We prove Properties 1 and 2. Properties 3 and 4 easily follow by construction and by definition of h-compatibility.

Proof of Property 1. We distinguish the following cases:

1. $D_p(\rho) < h$ and $N_\emptyset(\rho) < h$. Since $(\rho, \rho') \in R(h)$ and $h \in [1, n]$, by construction we easily deduce that $\rho = \rho'$.
2. $D_p(\rho) \geq h$. Since $(\rho, \rho') \in R(h)$, $D_p(\rho') \geq h$, $N_p(\rho') = N_p(\rho) = 0$, and either $N_\emptyset(\rho') = N_\emptyset(\rho)$, or $N_\emptyset(\rho) \geq h$ and $N_\emptyset(\rho') \geq h$. In both cases, by construction it easily follows that for each proper prefix σ of ρ, there exists a proper prefix σ' of ρ' such that $(\sigma, \sigma') \in R(h-1) \subseteq R(\lfloor \frac{h}{2} \rfloor)$.
3. $D_p(\rho) < h$ and $N_\emptyset(\rho) \geq h$. Since $(\rho, \rho') \in R(h)$, we have that $D_p(\rho') = D_p(\rho)$ (hence $\text{lst}(\rho) = \text{lst}(\rho')$), $N_p(\rho') = N_p(\rho)$, and $N_\emptyset(\rho') \geq h$. Let σ be a proper prefix of ρ. We distinguish the following three subcases:
   - **a.** $N_\emptyset(\sigma) < \lfloor \frac{h}{2} \rfloor$. Since $N_\emptyset(\rho) \geq h$, we have that $D_p(\sigma) \geq \lfloor \frac{h}{2} \rfloor$ and $|\sigma| = N_\emptyset(\sigma)$. Since $N_\emptyset(\rho') \geq h$, by taking the proper prefix σ' of ρ' having length $N_\emptyset(\sigma)$, we obtain that $(\sigma, \sigma') \in R(\lfloor \frac{h}{2} \rfloor)$.
   - **b.** $N_\emptyset(\sigma) \geq \lfloor \frac{h}{2} \rfloor$ and $D_p(\sigma) \geq \lfloor \frac{h}{2} \rfloor$. By taking the prefix σ' of ρ' of length $\lfloor \frac{h}{2} \rfloor$, we get that $(\sigma, \sigma') \in R(\lfloor \frac{h}{2} \rfloor)$.
   - **c.** $N_\emptyset(\sigma) \geq \lfloor \frac{h}{2} \rfloor$ and $D_p(\sigma) < \lfloor \frac{h}{2} \rfloor$. Since $\text{lst}(\rho) = \text{lst}(\rho')$, $N_p(\rho') = N_p(\rho)$, and $N_\emptyset(\rho') \geq h$, there exists a proper prefix σ' of ρ' such that $\text{lst}(\sigma') = \text{lst}(\sigma)$, $N_p(\sigma') = N_p(\sigma)$, and $N_\emptyset(\sigma') \geq \lfloor \frac{h}{2} \rfloor$. Hence $(\sigma, \sigma') \in R(\lfloor \frac{h}{2} \rfloor)$.

Thus in all the cases Property 1 holds.

Proof of Property 2. Let $(\rho, \rho') \in R(h)$ and σ be a non-empty track such that $\rho \cdot \sigma$ is a track. We distinguish the following cases:

1. $D_p(\rho) < h$. Since $(\rho, \rho') \in R(h)$, we have that $D_p(\rho') = D_p(\rho)$, $N_p(\rho) = N_p(\rho')$, and either $N_\emptyset(\rho') = N_\emptyset(\rho)$, or $N_\emptyset(\rho) \geq h$ and $N_\emptyset(\rho') \geq h$. Hence $\text{lst}(\rho) = \text{lst}(\rho')$ and by taking $\sigma' = \sigma$, we obtain that $(\rho \cdot \sigma, \rho' \cdot \sigma') \in R(h) \subseteq R(\lfloor \frac{h}{2} \rfloor)$.
2. $D_p(\rho) \geq h$ and $D_p(\sigma) < \lfloor \frac{h}{2} \rfloor$. It follows that $N_\emptyset(\rho \cdot \sigma) \geq \lfloor \frac{h}{2} \rfloor$. Since $D_p(\rho') \geq h$, there exists a track of the form $\rho' \cdot \sigma'$ such that $D_p(\rho' \cdot \sigma') = D_p(\rho \cdot \sigma)$, $N_p(\rho' \cdot \sigma') = N_p(\rho \cdot \sigma)$, and $N_\emptyset(\rho' \cdot \sigma') \geq \lfloor \frac{h}{2} \rfloor$. Hence $(\rho \cdot \sigma, \rho' \cdot \sigma') \in R(\lfloor \frac{h}{2} \rfloor)$.
3. $D_p(\rho) \geq h$ and $D_p(\sigma) \geq \lfloor \frac{h}{2} \rfloor$. Thus $D_p(\rho') \geq h$. If $N_\emptyset(\rho \cdot \sigma) < \lfloor \frac{h}{2} \rfloor$, then $N_\emptyset(\rho) = N_\emptyset(\rho')$. Therefore there exists a track of the form $\rho' \cdot \sigma'$ such that $N_\emptyset(\rho' \cdot \sigma') = N_\emptyset(\rho \cdot \sigma)$ and $D_p(\sigma') \geq \lfloor \frac{h}{2} \rfloor$. Otherwise $N_\emptyset(\rho \cdot \sigma) \geq \lfloor \frac{h}{2} \rfloor$ and there exists a track of the form $\rho' \cdot \sigma'$ such that $N_\emptyset(\rho' \cdot \sigma') \geq \lfloor \frac{h}{2} \rfloor$ and $D_p(\sigma') = \lfloor \frac{h}{2} \rfloor$. In both cases, $(\rho \cdot \sigma, \rho' \cdot \sigma') \in R(\lfloor \frac{h}{2} \rfloor)$.

Thus Property 2 holds, concluding the proof of the lemma. ◀

By applying Lemma 33 we deduce the following corollary which, together with Lemma 32, provides a proof of Lemma 21 (recall that $\mathcal{K}_n$ and $\mathcal{M}_n$ differ only in the initial state, which is s_0 for $\mathcal{K}_n$ and s_1 for $\mathcal{M}_n$).

▶ **Corollary 34.** *Let ψ be a balanced HS formula such that $|\psi| \leq n$ and $(\rho, \rho') \in R(|\psi|)$. Then $\mathcal{K}_n, \rho \models \psi$ iff $\mathcal{K}_n, \rho' \models \psi$.*

Proof. The proof is by induction on $|\psi|$. The cases for the Boolean connectives directly follow from the inductive hypothesis and the fact that $R(h) \subseteq R(k)$ for all $h, k \in [1, n]$ such that $h \geq k$. For the other cases, we proceed as follows:

- $\psi = p$: since $(\rho, \rho') \in R(1)$, ρ visits a state where p does not hold iff ρ' visits a state where p does not hold. Hence the result follows.
- $\psi = \langle B \rangle \theta$ (resp., $\psi = \langle \overline{B} \rangle \theta$): since ψ is balanced, θ is of the form $\theta = \theta_1 \wedge \theta_2$, where $|\theta_1| = |\theta_2|$. Hence $|\theta_1|, |\theta_2| \leq \lfloor \frac{|\psi|}{2} \rfloor$. We focus on the case $\psi = \langle B \rangle \theta$ (being the case $\psi = \langle \overline{B} \rangle \theta$ similar). Since $R(h)$ is an equivalence relation for each $h \in [1, n]$, it suffices to show that $\mathcal{K}_n, \rho \models \psi$ implies $\mathcal{K}_n, \rho' \models \psi$. Let $\mathcal{K}_n, \rho \models \psi$. Hence there exists a proper prefix σ of ρ such that $\mathcal{K}_n, \sigma \models \theta_i$ for $i = 1, 2$. Since $(\rho, \rho') \in R(|\psi|)$, by applying Lemma 33(1), there exists a proper prefix σ' of ρ' such that $(\sigma, \sigma') \in R(\lfloor \frac{|\psi|}{2} \rfloor)$. Since $R(\lfloor \frac{|\psi|}{2} \rfloor) \subseteq R(|\theta_i|)$ for $i = 1, 2$, by applying the inductive hypothesis we get that $\mathcal{K}_n, \sigma' \models \theta_i$ for $i = 1, 2$. Hence $\mathcal{K}_n, \rho' \models \psi$.
- $\psi = \langle E \rangle \theta$ (resp., $\psi = \langle \overline{E} \rangle \theta$): we apply Lemma 33(3) (resp., Lemma 33(4)) and the inductive hypothesis. ◀