Monitoring Second-Order Hyperproperties

Raven Beutner  
CISPA Helmholtz Center for  
Information Security  
Germany

Hadar Frenkel  
CISPA Helmholtz Center for  
Information Security  
Germany

Bernd Finkbeiner  
CISPA Helmholtz Center for  
Information Security  
Germany

Niklas Metzger  
CISPA Helmholtz Center for  
Information Security  
Germany

ABSTRACT

Hyperproperties express the relationship between multiple executions of a system. This is needed in many AI-related fields, such as knowledge representation and planning, to capture system properties related to knowledge, information flow, and privacy. In this paper, we study the monitoring of complex hyperproperties at runtime. Previous work in this area has either focused on the simpler problem of monitoring trace properties (which are sets of traces, while hyperproperties are sets of sets of traces) or on monitoring first-order hyperproperties, which are expressible in temporal logics with first-order quantification over traces, such as HyperLTL. We present the first monitoring algorithm for the much more expressive class of second-order hyperproperties. Second-order hyperproperties include system properties like common knowledge, which cannot be expressed in first-order logics like HyperLTL.

We introduce Hyper$^2$LTL$_f$, a temporal logic over finite traces that allows for second-order quantification over sets of traces. We study the monitoring problem in two fundamental execution models: (1) the parallel model, where a fixed number of traces is monitored in parallel, and (2) the sequential model, where an unbounded number of traces is observed sequentially, one trace after the other. For the parallel model, we show that the monitoring of the second-order hyperproperties of Hyper$^2$LTL$_f$ can be reduced to monitoring first-order hyperproperties. For the sequential model, we present a monitoring algorithm that handles second-order quantification efficiently, exploiting optimizations based on the monotonicity of subformulas, graph-based storing of executions, and fixpoint hashing. We present experimental results from a range of benchmarks, including examples from common knowledge and planning.

KEYWORDS

Runtime Verification, Common Knowledge, Multi-agent Systems

1 INTRODUCTION

Monitoring is a practical and scalable method for ensuring that the behavior of a complex system satisfies its formal specification. Unlike traditional verification techniques, such as model checking and theorem proving, monitoring does not analyze a system model, but instead works directly with the traces of the running system. Monitoring has been well studied for various types of trace properties, such as invariants or other properties expressible in standard temporal logics like linear-time temporal logic (LTL) [3, 15, 24]. By contrast, the study of monitoring algorithms for hyperproperties is still in an early stage.

Hyperproperties express relationships between multiple traces. Many properties related to knowledge, information flow, and privacy are hyperproperties and can, therefore, not be analyzed with methods for trace properties. The conceptual challenge is that while trace properties are sets of traces, hyperproperties are sets of sets of traces. So while the monitor of a trace property simply decides whether the observed trace is an element of the given set, the monitor of a hyperproperty must consider the entire set of traces produced by the system under observation (including any traces that will only be produced in the future) and decide whether this full set of traces is an element of the specified set of sets of traces.

Previous work on monitoring hyperproperties [1, 7, 8, 14] has focused on the special case of first-order hyperproperties, in particular on properties that can be expressed in the temporal logic HyperLTL [10]. HyperLTL extends LTL with first-order quantification over traces. First-order hyperproperties include many information-flow policies, such as noninterference and some simple notions of knowledge. Consider, for example, the property that an agent $i$ in a multi-agent system (MAS) knows that an LTL formula $\varphi$ is true whenever formula $\varphi$ is actually true. Knowing that $\varphi$ holds on a trace $\pi$ means that $\varphi$ must hold on all traces $\pi'$ that are indistinguishable from $\pi$ for agent $i$. The property is, therefore, expressed by the HyperLTL formula

$$\forall \pi. \forall \pi', \varphi[\pi] \implies (\pi \sim_i \pi' \implies \varphi[\pi']),$$

where $\pi \sim_i \pi'$ denotes that agent $i$ cannot distinguish between $\pi$ and $\pi'$, and $\varphi[\pi]$ encodes that $\varphi$ holds on trace $\pi$. A HyperLTL monitor for this formula detects a violation as soon as a trace $\pi'$ is observed that is indistinguishable from some previously seen trace $\pi$ and yet $\pi'$ does not satisfy $\varphi$.

Many hyperproperties of interest, however, cannot be expressed with first-order quantifiers alone. A prominent example is common
knowledge [11, 23, 32]. Common knowledge (CK) requires that every agent $i$ not only knows that $\varphi$ is true, but, additionally, knows that every other agent knows that agent $i$ knows that $\varphi$ is true. The fact that every agent knows this, must, in turn, be known by every other agent, and so on. Formally, common knowledge refers to a set of traces that is closed under the indistinguishability relations of all agents, and requires that $\varphi$ holds on all traces in this set. Hyper$^2$LTL [6] extends HyperLTL with second-order quantification over sets of traces and can thus express properties like CK. Since model-checking Hyper$^2$LTL formulas is undecidable, the verification method for Hyper$^2$LTL is based on approximation techniques. The question arises whether or not monitoring second-order hyperproperties suffers from a similar limitation. In this paper, we show that while the precise answer depends on the underlying execution model, for a large class of second-order hyperproperties the monitoring problem can, in fact, be solved effectively.

As a specification logic for monitoring second-order hyperproperties in MASs, we introduce the temporal logic Hyper$^2$LTL$_f$. The main difference to Hyper$^2$LTL is that monitoring deals with finite rather than infinite traces; Hyper$^2$LTL$_f$ therefore has a finite-trace semantics. Hyper$^2$LTL$_f$ furthermore includes past-time temporal operators and allows for arbitrary nestings of temporal operators and first/second-order quantifiers. Common knowledge in a MAS with agents $1, \ldots, n$ can be expressed as the Hyper$^2$LTL$_f$ formula:

$$∀\pi. \varphi[\pi] → ∃X. \pi ∈ X \land (∀\pi_1 ∈ X. ∀\pi_2. (∃_i \pi_1 ≈_i \pi_2 → \pi_2 ∈ X) ∧ ∀\pi' ∈ X. \varphi[\pi']).$$

The second-order quantifier $∃X$ specifies the existence of a set $X$ of traces so that $\pi$ is in $X$, and all traces that are indistinguishable from some trace in $X$ for some agent are also in $X$. Given this set, a monitor detects a violation by finding a trace in the set that does not satisfy $\varphi$.

Two fundamental execution models in the monitoring of hyperproperties are the parallel and sequential models (cf. [14]): in the parallel model, a fixed number of traces is monitored in parallel; in the sequential model, an unbounded number of traces is observed sequentially, one trace after the other.

For the parallel model, we show that the monitoring of the second-order hyperproperties expressed in Hyper$^2$LTL$_f$ can be reduced to monitoring first-order hyperproperties. For the sequential model, we show that the monitoring problem is undecidable in general, but becomes feasible for the practically relevant class of monotone second-order hyperproperties. A second-order hyperproperty is $⊕$-monotone if its satisfaction on some set of traces implies that it is also satisfied on any superset of this set. Conversely, a second-order hyperproperty is $⊗$-monotone if a violation on some set implies its violation on any superset. Monotonicity thus allows the monitor to provide definitive results that hold irrespectively of the traces that may still arrive in the future.

We introduce an inference system for monotonicity and present a monitoring algorithm that iteratively checks the given set of traces until it can produce a decisive answer. The algorithm has been implemented in a tool called MoSo. We report on encouraging evaluation results for MoSo on several benchmarks, including examples from common knowledge and planning.

Related Work. Giacomo and Vardi [19] first introduced a logic for interpreting LTL on finite traces, called LTL$_f$, specifically designed for AI systems. Since then, LTL$_f$ and its variants have been used, e.g., for model checking [2, 30], satisfiability analysis [16], synthesis [20, 21], and planning [9]. Knowledge in combination with LTL dates back to Fagin et al. [11] and has found many applications in MASs [12, 25, 26]. Logics for hyperproperties have mostly been obtained by extending temporal logics with explicit quantification over traces or paths, such as HyperLTL [10], HyperQPTL [29], HyperPDL [22], HyperATL* [4, 5], and HyperLDL$_f$ [18]. Hyper$^2$LTL adds second-order quantification over sets of traces [6]. Tools for model-checking knowledge in multi-agent systems include MCK [17] and MCMAS [27], the latter was also extended to finite trace semantics [26]. While these approaches and tools implement solutions explicitly for knowledge, Hyper$^2$LTL$_f$ generalizes to more general second-order hyperproperties.

2 Hyper$^2$LTL$_f$

We consider a finite set of atomic propositions $AP$ and define $Σ := 2^{AP}$. We define Hyper$^2$LTL$_f$ as a finite-trace extension of Hyper$^2$LTL [6]. Let $V = \{π_1, π_2, π_3, \ldots\}$ be a set of (first-order) trace variables and $W = \{X, Y, \ldots\}$ a set of (second-order) set variables. We assume a special second-order variable $sys ∈ W$ that we use to refer to the set of all system traces. Hyper$^2$LTL$_f$ formulas are defined by the following grammar:

$$\varphi := a_π | ¬ \varphi | \varphi ∧ \varphi | Q \varphi | ⊕ \varphi | ⊗ \varphi | U \varphi | F \varphi | ⋄ \varphi | S \varphi | Q π ∈ X. \varphi | Q X. \varphi$$

where $a ∈ AP, π ∈ V, X ∈ W$, and $Q ∈ \{∀, ∃\}$ is a quantifier. In Hyper$^2$LTL$_f$, we have the future temporal operators (strong) $next$ and $until$ $U$, as well as their past counterparts $previously$ $⊗$ and $since$ $S$. We use the derived Boolean constants $true, false$, and connectives $∧, →, ↔$, and the temporal operators eventually $⋄ \varphi = true U \varphi$, once $□ \varphi = true S \varphi$, globally $□ \varphi = ¬ □ ¬ \varphi$, and historically $□ □ \varphi = ¬ □ ⊤ ¬ \varphi$.

Semantics. The semantics of Hyper$^2$LTL$_f$ is defined with respect to a trace length $m ∈ N$ and a set of traces $T ⊆ 2^m$. As for HyperLTL, we use a trace assignment $π : V → 2^m$ mapping trace variables in $V$ to finite traces of length $m$. Additionally, we maintain a second-order assignment $Δ : W → 2^{2^m}$ mapping variables to sets of finite traces of length $m$. The semantics is then as follows:

$$\begin{align*}
Π, Δ, i ⊨ T a_π & \quad \text{iff} \quad a ∈ Π(π)(i) \\
Π, Δ, i ⊨ T ¬ \varphi & \quad \text{iff} \quad Π, Δ, i ⊭ T \varphi \\
Π, Δ, i ⊨ T φ_1 ∧ φ_2 & \quad \text{iff} \quad Π, Δ, i ⊨ T φ_1 \land Π, Δ, i ⊨ T φ_2 \\
Π, Δ, i ⊨ T Q \varphi & \quad \text{iff} \quad i < m − 1 \land Π, Δ, i + 1 ⊨ T \varphi \\
Π, Δ, i ⊨ T ⊗ \varphi & \quad \text{iff} \quad i > 0 \land Π, Δ, i − 1 ⊨ T \varphi \\
Π, Δ, i ⊨ T φ_1 U φ_2 & \quad \text{iff} \quad \exists j ≤ m. Π, Δ, j ⊨ T φ_2 \land \forall k ≤ j < i. Π, Δ, k ⊨ T φ_1 \\
Π, Δ, i ⊨ T φ_1 S φ_2 & \quad \text{iff} \quad \exists j ≤ i. Π, Δ, j ⊨ T φ_2 \land \forall k ≥ j > i. Π, Δ, k ⊨ T φ_1
\end{align*}$$

1We restrict to traces of the same length to obtain simpler semantics. Without this restriction, we would have to deal with combinations of traces with different lengths, making traversal difficult. In practice, we can either pad traces to the length of the longest trace or crop them to the length of the shortest trace (cf. [13, 14]).
The temporal and boolean operators are evaluated as expected. Atomic formula \( a_r \) holds in step \( i \) if \( a \) holds in the \( i \)th step on the trace bound to \( \pi \). When quantifying over a trace \( \pi \in X \), we quantify over all traces in its current second-order model (as given by \( \Delta \)), and when quantifying over a set of traces, we consider all possible subsets of \( T \). A set of traces \( T \subseteq \Sigma^\omega \) satisfies \( \varphi \), written \( T \models \varphi \), if \( \emptyset, \{ \} \models \varphi \). Note that we bind the second-order variable \( \pi \) to \( T \); by writing \( \forall \pi \in \Sigma \) and \( \exists \pi \in \Sigma \) we can thus quantify over the traces in \( T \) (as in HyperLTL).

HyperLTL and Hyper\(^\omega\)LTL\(_f\). There are four key differences between the logics Hyper\(^\omega\)LTL\(_f\) and HyperLTL. Firstly, Hyper\(^\omega\)LTL\(_f\) adds past operators, which allow for exponentially more succinct definitions of properties [28] and offers a convenient syntax to specify knowledge properties (cf. Example 1). Secondly, the semantics of Hyper\(^\omega\)LTL\(_f\) is defined w.r.t. finite traces, which occur much more frequently in MAS-related domains [2, 11, 19, 20]. Thirdly, Hyper\(^\omega\)LTL\(_f\) allows quantification under temporal operators. In HyperLTL [10] or Hyper\(^\omega\)LTL [6], a formula consists of a quantifier prefix followed by a quantifier-free LTL body. In contrast, in Hyper\(^\omega\)LTL\(_f\), we can interchange temporal operators and quantification. For example, \( \forall \pi \in \Sigma \bigcirc \forall \pi' \in \Sigma \bigcirc a_r \leftrightarrow a_{r'} \) states that on any trace, we can find a timepoint from which point on all other traces agree in \( a \). Note that despite interleaving temporal operators and quantification, Hyper\(^\omega\)LTL\(_f\) has a *linear-time semantics* (different from, e.g., HyperCTL\(^*\) [10]), which is crucial in our monitoring setting. And lastly, and most importantly, Hyper\(^\omega\)LTL\(_f\) features second-order quantification over sets of traces. Note that while second-order quantification in Hyper\(^\omega\)LTL [6] ranges over arbitrary sets of traces, Hyper\(^\omega\)LTL\(_f\) quantifies over subsets of \( T \).

**Example 1 (Eventual Knowledge).** As a running example, we use the sender-receiver MAS modeled by transition system in Figure 1, which is a variation of an example presented in [32]. In this system, agent 1 attempts to send a message until it is either received by agent 2, or delayed for one step and received afterward. The two agents each only have a partial view of the system: agent 1 observes the sending state (i.e., only observes AP \( s \)) and thus cannot distinguish if the message is immediately received or delayed for one step. Likewise, agent 2 only observes if the message was received (modeled by AP\(_R\)). We want to express that whenever a message is received twice (i.e., two consecutive \( r \)s occur), agent 1 eventually knows that it was received. As in epistemic logics, an agent knows a property on a trace if it holds on all indistinguishable traces, which we can express as follows:

\[
\forall \pi \in \Sigma, (\Diamond (r_\pi \land \bigcirc r_{\pi'})) \implies (\forall \pi' \in \Sigma, \pi \sim_1 \pi' \implies \Diamond r_{\pi'}),
\]

where we define \( \pi \sim_1 \pi' := \Box \pi' \leftrightarrow \pi_{\sigma_x} \) to state that agent 1 cannot distinguish \( \pi \) and \( \pi' \) up to the current time step. Here, we quantify over all traces \( \pi \) that visit the receiving state twice. For any such \( \pi \), we require that eventually, all traces \( \pi' \) that are \( \sim_1 \) to \( \pi \) up to the current time point \( \sigma_x \) are indistinguishable from \( \pi \), satisfy \( \Diamond r_{\pi'} \). This property holds on the system in Figure 1. For example, consider the trace \( s^0rr \), where sending takes \( n \) steps and immediately afterward the message is received. In the last time step, this trace is indistinguishable (for agent 1) from \( s^0rr \) and \( s^0dr \), and both of these traces satisfy \( \Diamond r \).
Example 2 (Common Knowledge and Fixpoints). We illustrate our fixpoint construct by continuing on Example 1. Assume we do not want to state that on all traces where \( \diamond (r \land \Box r) \) holds, agent 1 knows that \( \Diamond r \) holds (cf. Example 1), but rather that it is common knowledge in the group of agents 1 and 2 that \( \Diamond r \) holds. As argued in Section 1, this requires us to iteratively compute the set of all traces that cannot be distinguished by some agent. While the formulation of CK in Section 1 expressed this using full second-order quantification, we can observe that the set we are interested in is actually defined by a fixpoint:

\[
\forall \pi \in \text{sys.} (\Box (r \land \Box r \pi)) \rightarrow \Diamond (\text{fix}(X, \xi_1, \xi_2) \pi' X. \Diamond r \pi'),
\]

where \( \xi_1 := \text{true} \pi \in X \), and \( \xi_2 \) is defined as

\[
\forall \pi_1 \pi_2 \pi_3 \in \text{sys.} (\xi_1 \land \xi_2 \pi_1 \pi_2) \rightarrow \pi_2 \pi \in X.
\]

Here, we define \( \pi \sim_2 \pi' := \Box (r \leftarrow r \pi) \), similar to \( \sim_1 \) in Example 1. For each \( \pi \), we specify a fixpoint-defined set \( X \). This set \( X \) includes \( \pi \) (constraint \( \xi_1 \)), and whenever we find some \( \pi_1 \pi_2 \pi_3 \in X \) that some agent cannot distinguish from some \( \pi_3 \pi_2 \pi_3 \pi_2 \pi_3 \), we add \( \pi_2 \pi_2 \pi_3 \) to \( X \) (constraint \( \xi_2 \)). Then, we state that all traces \( \pi' \) in \( X \) should satisfy \( \Diamond r \). The above formula does not hold on the MAS modeled in Figure 1. Consider the trace \( \pi = sr^n \). It is easy to see that

\[
sr^n \sim_1 sdr^{n-1} \sim s^n r \sim s^n d.
\]

Consequently, the trace \( s^n d \) will be included in \( X \), disproving that \( \pi \) is common knowledge on \( \pi \).

3 MONITORABILITY

We are interested in monitoring a Hyper\( ^2 \text{LTL} \) formula, i.e., we do not have access to the system as a whole but rather observe executions of the system and conclude whether the set of traces we have seen so far suffices to conclude the satisfaction or violation of the property. This is, unsurprisingly, not possible for all properties. Some properties are not monitorable, i.e., never allow a positive or negative answer from the monitor.

3.1 Monitorability of Trace Properties

Let us recall the concept of monitorability in the simpler setting of a trace property, where we observe one step of the trace at a time.

Definition 1. A trace property \( L \subseteq \Sigma^* \) is monitorable if

\[
\forall u \in \Sigma^*. \exists v \in \Sigma^* \forall w \in \Sigma^* : u w v \in L \lor (\forall w \in \Sigma^*: u w v \notin L).
\]

Intuitively, the definition states that whatever finite trace \( u \) we start from, some extension \( u w v \) of \( u \) allows the monitor to report the satisfaction or violation of \( L \). This can either be the case because all extensions of \( u w v \) satisfy \( L \) (the first disjunct) or all extensions violate \( L \) (the second disjunct). In this paper, we are not monitoring a trace property but a hyperproperty, i.e., we may need to observe multiple traces. Depending on the setting in which the monitor is deployed, there can be different modes in which those traces are presented to the monitor. We focus on the parallel model and the sequential model [14]. In the former, the number of traces is fixed \( a \text{ priori} \), and the monitor observes consecutive time steps one by one. In the latter model, the monitor observes the (finite) traces sequentially, increasing the cardinality of the trace set in every time step. We discuss both models and their impact on the monitorability of Hyper\( ^2 \text{LTL} \) properties in the following.

3.2 The Parallel Model

In the parallel model, the number of traces (executions) is fixed to some number \( b \in \mathbb{N} \), and each time step reveals an additional position on each of the traces. Similar to the monitorability of trace properties, the monitor thus gets presented more and more positions on the same \( b \) traces. We refer to [14] for more details.

As the parallel model fixes the number of sessions (traces), it turns out that second-order quantification does not yield any additional expressiveness. Intuitively, quantifying over a set of traces is equivalent to quantifying over some finite subset of at most \( b \) traces, which is expressible in Hyper\( ^2 \text{LTL} \) with first-order trace quantification. Formally, given bound \( b \in \mathbb{N} \) and a Hyper\( ^2 \text{LTL} \) formula \( \varphi \), we can unfold \( \varphi \) into a formula that uses no second-order quantification and is equivalent on any set of traces with at most \( b \) traces.

We maintain a partial function \( M : \{\forall, \exists\} \rightarrow \mathbb{N}^2 \) mapping second-order variables to a set of trace variables. We then recursively define the Hyper\( ^2 \text{LTL} \) formula \( [\varphi]_{b, M} \) as follows:

\[
[a \pi]_{b, M} := \sigma_a,
\]

\[
[b \varphi]_{b, M} := \bigwedge_{\pi \in X} [\varphi]_{b, M} \quad \text{for } \sigma \in (\neg, \odot, \odot),
\]

\[
[\forall \pi \varphi \odot]_{b, M} := \bigwedge_{\pi \in X} [\forall \pi \varphi \odot]_{b, M} \quad \text{for } \sigma \in \{\forall, \exists\}.
\]

\[
[\exists \pi \varphi]_{b, M} := \bigvee_{\pi \in X} [\exists \pi \varphi]_{b, M} \quad \text{for } \sigma \in \{\forall, \exists\}.
\]

\[
[\forall \pi \in X \varphi]_{b, M} := \bigwedge_{\pi' \in X} [\forall \pi \varphi]_{b, M} \quad \text{for } \sigma \in \{\forall, \exists\}.
\]

\[
[\exists \pi \in X \varphi]_{b, M} := \bigvee_{\pi' \in X} [\exists \pi \varphi]_{b, M} \quad \text{for } \sigma \in \{\forall, \exists\}.
\]

Here, \( \varphi[\pi' / \pi] \) denotes the formula in which all free occurrences of \( \pi \) are replaced with \( \pi' \). For quantifier-free formulas, we simply maintain the structure. Since we assume that there are at most \( b \) traces, we can replace second-order quantification over a set \( X \) with first-order quantification over \( b \) fresh (first-order) trace variables \( \pi_1, \ldots, \pi_b \). We record the first-order trace variables that we used to replace \( X \) within the auxiliary mapping \( M \). For first-order quantification \( \forall \pi \in X \) (resp. \( \exists \pi \in X \)), we then conjunctively (resp. disjunctively) consider all traces \( \pi' \in M(X) \) in place of \( \pi \). A simple induction shows:

Proposition 1. Let \( \varphi \) be any Hyper\( ^2 \text{LTL} \) formula and \( b \in \mathbb{N} \). For any set of traces \( T \) with \( |T| \leq b \) we have \( T \models \varphi \) iff \( T \models [\varphi]_{b, 0} \).

Assuming a fixed bound \( b \in \mathbb{N} \) (as in the parallel model), Proposition 1 states that monitoring a Hyper\( ^2 \text{LTL} \) formula \( \varphi \) is equivalent to monitoring \( [\varphi]_{b, 0} \). As \( [\varphi]_{b, 0} \) uses no second-order quantification, we can apply all the techniques and tools developed for first-order logics like Hyper\( ^2 \text{LTL} \) [14].

3.3 The Sequential Model

We now consider the sequential model, the main object of study in this paper. In this model, traces arrive one at a time and the number of traces is unbounded. Our model is similar to the sequential
model used in [14] for HyperLTL, with the exception that traces in our model are of finite (instead of infinite) length. This focus on finite traces makes our approach (and tool) applicable to many MAS-related settings, where executions are finite, and tasks are subsequent system executions.

**Definition 2.** A hyperproperty \( \mathcal{H} \subseteq 2^{\Sigma^*} \) is monitorable in the sequential model if

\[
\forall \mathcal{U} \subseteq \Sigma^*, \exists \mathcal{V} \subseteq \Sigma^*.
(\forall \pi \subseteq \Sigma^*, (U \cup V \cup W) \in H) \lor (\forall \pi \subseteq \Sigma^*, (U \cup V \cup W) \not\in H).
\]

That is, for any \( U \) there exists some \( V \) such that \( U \cup V \) allows some definitive answer by the monitor. This can either be the case because all extensions of \( U \cup V \) satisfy \( H \) (the first disjunct) or all extensions violate \( H \) (the second disjunct).

**Example 3.** Consider the CK formula in Example 2. The formula is monitorable in the sequential model: We can always add traces to ensure that CK does not hold, i.e., add some indistinguishable trace that does not satisfy \( \square r \). No matter what additional traces are observed, CK remains violated. \( \triangle \)

**Theorem 1.** Deciding if a Hyper\(^2\)LTL\(_f\) formula is monitorable in the sequential model is undecidable.

Proof. Finkbeiner et al. [14] showed that monitorability is undecidable for HyperLTL with finite traces in the sequential model. Since HyperLTL with finite-trace semantics is a strict subset of Hyper\(^2\)LTL\(_f\), monitorability of Hyper\(^2\)LTL\(_f\) is undecidable. \( \square \)

4 MONOTONICITY

In the following, we focus on the sequential model. Theorem 1 rules out the possibility of an algorithm that monitors all possible Hyper\(^2\)LTL\(_f\) formulas. Instead, we focus on fragments of formulas for which we can provide definitive answers. Given a finite set of traces, we want to provide a monitoring result irrespective of what traces will arrive in the future. To ensure this, the truth value of the formula may not change when additional traces arrive; that is, it should be monotone. We distinguish between \( \oplus \)-monotonicity and \( \ominus \)-monotonicity. In \( \oplus \)-monotonicity, any model that satisfies the formula will continue to satisfy it no matter what traces are added. In \( \ominus \)-monotonicity, any model that violates the formula will continue to violate it no matter what traces are added.

**Definition 3.** A Hyper\(^2\)LTL\(_f\) formula \( \varphi \) is \( \oplus \)-monotone (resp. \( \ominus \)-monotone) if for any set of traces \( T \) such that \( T \models \varphi \) (resp. \( T \not\models \varphi \)), for any larger set \( T' \supseteq T \), we have \( T' \models \varphi \) (resp. \( T' \not\models \varphi \)).

We say that a formula is monotone if it is either \( \oplus \)-monotone or \( \ominus \)-monotone. Once we detect that a \( \oplus \)-monotone (resp. \( \ominus \)-monotone) formula holds (resp. is violated) on the set of traces seen so far, our monitor can conclude that \( \varphi \) holds (resp. does not hold) regardless of what traces will be observed in the future. The CK formula in Example 2 is \( \ominus \)-monotone, i.e., whenever the current set of traces violates CK, no additional set of traces can change it. In many cases, monotonicity implies monitorability:

**Proposition 2.** Let \( \varphi \) be a Hyper\(^2\)LTL\(_f\) formula. If one of the following holds, then \( \varphi \) is monitorable.

\[
\begin{align*}
\Gamma \vdash a_\pi : \emptyset & \quad \Gamma \vdash a_\pi : \emptyset & \quad \Gamma \vdash \varphi : \emptyset & \quad \Gamma \vdash \varphi : \emptyset & \quad \Gamma \vdash a_\pi : \emptyset \\
\circ \in \{\lor, \land\} & \quad \Gamma \vdash \varphi : \alpha & \quad \Gamma \vdash \varphi : \emptyset & \quad \Gamma \vdash \varphi : \emptyset & \quad \Gamma \vdash \varphi : \emptyset \\
\Gamma \vdash \Diamond \varphi : \alpha & \quad \Gamma \vdash \varphi : \emptyset & \quad \Gamma \vdash \Diamond \varphi : \emptyset & \quad \Gamma \vdash \varphi : \emptyset & \quad \Gamma \vdash \varphi : \emptyset \\
X \in \Gamma & \quad \Gamma \vdash \exists \alpha \in X. \varphi : \emptyset & \quad \Gamma \vdash \forall \alpha \in X. \varphi : \emptyset & \quad \Gamma \vdash \varphi : \emptyset & \quad \Gamma \vdash \Diamond \varphi : \emptyset \\
\Gamma \cup \{X\} \vdash \varphi : \emptyset & \quad \Gamma \vdash \varphi : \emptyset & \quad \Gamma \vdash \Diamond \varphi : \emptyset & \quad \Gamma \vdash \Diamond \varphi : \emptyset & \quad \Gamma \vdash \Diamond \varphi : \emptyset
\end{align*}
\]

**Figure 2: Inference system for monotonicity**

(1) \( \varphi \) is \( \oplus \)-monotone and has at least one finite model (i.e., a finite set of traces \( T \) such that \( T \models \varphi \)).

(2) \( \varphi \) is \( \ominus \)-monotone and \( \neg \varphi \) has at least one finite model.

The proof directly follows from the definitions of monotonicity (Definition 3) and monitorability (Definition 2). The converse statement does not hold, even for the first-order fragment of Hyper\(^2\)LTL\(_f\).

**Example 4.** Consider the formula

\[
\varphi := \forall \pi. \exists \alpha' . (\alpha' \neq \alpha \land [a_\pi \leftrightarrow a_{\alpha'}]) \lor \exists \psi' . \Diamond (b_{\psi'})
\]

and the trace sets \( T := \{a, c\}^m, \{a\}^m \), \( T' := T \cup \{c\}^m \), and \( T'' := T' \cup \{b\}^m \). It is easy to see that \( T \models \varphi, T' \not\models \varphi, \) and \( T'' \models \varphi \). Formula \( \varphi \) thus cannot be monotone. Yet, \( \varphi \) is monitorable: for every set \( U \), we can add \( V := \{b\}^m \), and every extension of \( U \cup V \) satisfies \( \varphi \) (cf. Definition 2).

An Inference System for Monotonicity. In our monitoring algorithm, we use monotonicity to provide definitive monitoring answers. To statically determine the monotonicity of (sub)formulas, we use a deductive (type-like) inference system. The judgments in our system are of the form \( \Gamma \vdash \varphi : \alpha \), where \( \alpha \in \{\emptyset, \ominus\} \), and \( \Gamma \models \{X_1, \ldots, X_n\} \) is a context assumption. Intuitively, \( \Gamma \models X \) assumes that \( X \) has a unique model that only grows when more traces arrive (i.e., the model is monotonically increasing); and \( \Gamma \vdash \varphi : \ominus \alpha \) implies that, under the context assumptions in \( \Gamma \), \( \varphi \) is \( \oplus \)-monotone (resp. \( \ominus \)-monotone). The inference rules of our system are depicted in Figure 2. Most rules in our system are straightforward: atomic propositions are both \( \oplus \)-monotone and \( \ominus \)-monotone, temporal operators preserve monotonicity, and negation "flips" monotonicity. More interesting are the rules for quantification. A formula \( \varphi := \exists \pi \in X. \varphi' \) is \( \ominus \)-monotone if \( \varphi' \) is \( \ominus \)-monotone and \( X \in \Gamma \) (i.e., the model of \( X \) only increases); assume that \( \varphi \) holds on the current set of traces \( T \), i.e., there exists a trace \( t \in X \) that is a witness for the satisfaction of \( \varphi' \). As \( X \in \Gamma \), the model of \( X \) can only grow when more traces are added to \( T \), so we can always reuse \( t \) as a witness. Likewise, models of a fixpoint-defined second-order variable \( X \) only grow larger with the arrival of new traces, so we add \( X \) to \( T \). For full second-order quantification, we can make no assumption on how model(s) behave when future traces arrive, so we add no assumptions. In our inference system, we initially set the
context to be Γ = {sys}, as the set of all traces (bound to the special variable sys ∈ W) only grows larger. An easy induction shows:

**Proposition 3.** Assume (sys) ⊓ φ : (resp. (sys) ⊔ φ : ω), then φ is ⊓-monotone (resp. ⊔-monotone).

Note that our system is not complete, i.e., the converse of Proposition 3 does not hold. It is easy to see that our system allows us to conclude that the CK property in Example 2 is ⊓-monotone.

## 5 Monitoring Algorithm

We depict our basic monitor for a formula φ in Algorithm 1. As a first step, we (1) compute a monotonicity map monMap using the inference system from Figure 2. Here, monMap maps each subformula φ’ of φ to monMap(φ’) ∈ 2[φ, ω] indicating whether φ’ is in a given context Π, Δ, i − is monotone; and (2) initialize the set T of system traces to the empty set. During the execution of the monitor, whenever a new trace t arrives, we add it to T and check if T satisfies φ by calling function check (which we discuss in Section 5.1). As we argued in Section 4, our monitor can only provide a definitive SAT answer in case φ is Π-monotone and the current traces satisfy φ, and UNSAT in case it is Π-monotone and the current set of traces violates φ. Otherwise, we await further traces and repeat.

### 5.1 Incremental Model Checking

The core of our monitoring algorithm lies in our recursive model-checking function check given in Algorithm 2. On a high level, check casts the semantics of Hyper²LTL₁ into an executable program. For now, we ignore the program fragments marked with a gray background; these concern the hashing-based optimizations we will discuss in Section 5.3. In line 6, we match on the structure of φ. We only include a selection of cases (the others can be inferred easily). In case φ is an AP, we can immediately decide its truth value (line 7); for the cases of negation, conjunction, and temporal operators we perform the expected recursive call(s). For first-order quantification (line 9), we iteratively check all traces assigned by Δ. For full second-order quantification (line 16), we check all subsets of T. In the case of fixpoint-based second-order quantification (line 21), we compute the (unique) solution to the fixpoint definition (line 23) by calling computeFix (discussed in Section 5.2).

### 5.2 Fixpoint Computation

As we argue in Section 2, fixpoints are expressive enough for most practical properties and are much easier to handle algorithmically. Given a set T with n traces, we can compute the fixpoint in polynomial time solution using Knaster-Tarski fixpoint iteration [31] instead of trying all possible 2ⁿ subsets of T (as needed for full second-order quantification). The input to Algorithm 3 is the fixpoint formula and the current set A of traces in the (to-be) fixpoint set. Initially, A = ∅. We check if A satisfies the fixpoint constraints ξ₁, ..., ξₖ: if we find traces t₁ ∈ Δ(X₁), ..., tₙ ∈ Δ(Xₙ) that satisfy the step constraint (which we check via a mutually recursive call in line 6), we add the trace required by the fixpoint constraint to A and repeat via a recursive call (line 7).

### 5.3 Monitoring Optimizations

When implemented without optimizations, Algorithm 2 checks the Hyper²LTL₁ formula on the current set of traces, re-performing this check whenever a new trace arrives. This is often infeasible in practice, as with each new trace, the number of traces and, therefore, the verification time increases. Instead, whenever we add a trace t to the current set of traces T, we want to reuse as much of the computation we have already performed on T in previous steps. We discuss three areas where such re-use is possible: hashing of verification results for subformulas, hashing of fixpoint results, and hashing of witnesses.

**SAT Hashing.** Our first optimization is the hashing of verification results of the check for subformulas. If φ is a given evaluation context (Π, Δ, i) a Π-monotone (resp. Π-monotone) subformula φ was satisfied (resp. violated) in a previous iteration, then φ remains satisfied (resp. violated) in context (Π, Δ, i), even if more traces are added to T. We therefore use a (hashing) function hash (initially set to the function with empty domain) and check if we have hashed (Π, Δ, i, φ) (Algorithm 2, lines 2 and 4). Here, we, e.g.,
We have implemented our monitoring algorithm and the optimizations from Section 5.3 in a tool called MoSo. In this section, we demonstrate that MoSo can monitor complex second-order hyper-properties that are out of reach of existing monitoring and model-checking frameworks.

### 6.1 Running Example

The first experiment is the scalability of MoSo in the length of the traces. We verify the formula in Example 2 against traces of the running example in Figure 1. We generate traces for each instance of lengths 20, 30, 40, ..., 80 and measure the time it takes to find the violation for the starting trace $s_0^m$ in Table 1. The monitor correctly concludes in all runs that the common knowledge formula is violated.

<table>
<thead>
<tr>
<th># children</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t_{\text{noFix}}$</td>
<td>0.1</td>
<td>1.1</td>
<td>TO</td>
<td>TO</td>
<td>TO</td>
<td>TO</td>
<td>TO</td>
<td>TO</td>
</tr>
</tbody>
</table>

### 6.2 Common Knowledge in MASs

We monitor common knowledge in a scalable instance of the muddy children’s puzzle [11]. The muddy children puzzle is a MAS between children $1, \ldots, n$, where a number of rounds of communication are used to achieve common knowledge about how many children are muddy. For each child (agent) $i$, we have an AP $m_i$ indicating if $i$ is muddy and an AP $d_i$ indicating if $i$ declared that it is muddy. We can then express that theuddiness of all children is CK after $b \in \mathbb{N}$ steps of communication (which we call the communication-bound) as the following Hyper$\text{LTL}_f$ formula $\varphi_{\text{mud}}$:

$$\forall \pi \in \text{sys}. \bigwedge_{i=1}^n \text{fix}(X, \xi_1, \xi_2), \forall \pi_1 \in X, \forall \pi_2 \in X. \bigwedge_{a \in \{m_1, \ldots, m_n\}} (a_{\pi_1} \leftrightarrow a_{\pi_2}),$$

where $\xi_1 := \text{true} \rightarrow \pi \in X$, and $\xi_2$ is defined as

$$\forall \pi_1 \in X. \forall \pi_2 \in \text{sys}. \left( \bigwedge_{i=1}^n \left( \bigwedge_{a \in \{d_1, \ldots, d_n\} \cup \{m_{1}, \ldots, m_n\}} (a_{\pi_1} \leftrightarrow a_{\pi_2}) \right) \right) \rightarrow \pi_2 \in X.$$

Intuitively, the fixpoint captures all traces that some child $i$ cannot distinguish from $\pi$ in the first $b$ steps, i.e., we add all traces that agree on all APs except $m_i$. Formula $\varphi_{\text{mud}}$ then requires that all traces in $X$ agree on the mudliness of all children. The muddy children’s MAS violates $\varphi_{\text{mud}}$ if and only if $b < n$.

### Scalability

For the game with $n$ children, we sample random traces and use MoSo to monitor CK (we set the communication-bound to be $b = \lceil \frac{n}{2} \rceil$ so CK does not hold). We report the runtime in Table 2. We observe that even if the number of children grows, our monitor will find violations to common knowledge quickly and
thus runs very effectively. In contrast, model-checking of CK [6] is, currently, only possible up to \( n = 4 \). This supports our claim that monitoring second-order hyperproperties is a useful lightweight technique that scales in settings where full verification is infeasible.

**Fixpoints vs Second-Order Sets.** The muddy children’s puzzle also highlights the importance of fixpoint formulas compared to arbitrary second-order quantification. To test this empirically, we consider the same CK property but express it directly using full second-order quantification (similar to the formula in Section 1). We give the runtime of MoSo on the non-fixpoint-based formula in Table 2 (\( t_{\text{fix}} \)). When using general second-order quantification, MoSo already times out for 4 children (due to the exponential cost of considering all subsets of traces). This attests to the importance of fixpoints for scalable monitoring.

**Impact of Communication-bound.** Finding violations to the CK formula heavily depends on the communication-bound. With increasing communication-bound, the probability of sampling traces that violate CK decreases as more information is observed by each child. On average, we thus need to see more traces until our monitor reports a violation. We test this empirically by monitoring the muddy children puzzle with a varying number of children and communication-bound and report on the average number of traces (across 10 monitor runs) the monitor observes before concluding that CK does not hold. We depict the results in Table 3. A larger communication bound clearly increases the number of traces the monitor processes until a violation is reported.

### 6.3 Impact of Optimizations

Common knowledge also serves as a useful baseline to highlight the importance of our hashing-based optimizations. To demonstrate this, we check the following CK formula

\[
\exists \pi \in \text{sys}. \diamond \text{fix} (X, \text{true} \rightarrow \pi \in X, \forall \pi_1 \in X. \forall \pi_2 \in \text{sys}.
\]

\[
(\square (a_{\pi_1} \leftrightarrow a_{\pi_2}) \lor \Box (b_{\pi_1} \leftrightarrow b_{\pi_2}) \rightarrow \pi_2 \in X). \forall \pi' \in X. c_{\pi'},
\]

which states that on some path \( \pi \), \( c \) is eventually CK, given that one agent observes \( AP \) \( a \) and one agent observes \( AP \) \( b \). We use MoSo to check the above formula on randomly generated instances with a varying number of traces. We depict the results in Figure 3. We observe that all optimizations improve upon the baseline.

### 6.4 Planning Analysis

In our last experiment, we use our monitor to detect reachability in a graph based on the observation of random paths. For example, when observing traces \( t_1 = abc \), and \( t_2 = bbd \), we can conclude that the trace \( t = abd \) exists in the graphs, even without observing it. The inference of \( t \) can easily be stated as a fixpoint constraint: We differentiate between time-sensitive (t-sen.) and time-insensitive (t-ins.) paths. In the former model, we can only “combine” two paths if they visit the same state at the same time (step). In the latter model, we can even combine two paths if they visit the same state at potentially different times. We report on the time and number of observed traces it takes to find a correct plan in Table 4. As expected, the time-insensitive case terminates earlier than the time-sensitive case since more combinations are possible, increasing the chance of concluding the existence of a path from source to target.

### 7 CONCLUSION

We have presented Hyper\(^2\)\texttt{LTL}_f, the first logic to express second-order hyperproperties for finite trace models. Our experiments show that our monitoring algorithm scales to larger instances well beyond the reach of complete methods such as model checking. Monitoring of complex and important second-order hyperproperties (particularly in the MAS domain), is thus a viable middle ground to obtain rigorous guarantees while maintaining scalability.

**ACKNOWLEDGMENTS**

This work was supported by the European Research Council (ERC) Grant HYPER (101055412), by the German Research Foundation (DFG) as part of TRR 248 (389792660), and by the German Israeli Foundation (GIF) Grant No. I-1513-407.2019.
REFERENCES


