

A Complete Program Logic for Compositional Linearizability

Eashan Hatti ✉ 

Yale University, New Haven, CT, USA

Arthur Oliveira Vale ✉ 

Yale University, New Haven, CT, USA

Zhongye Wang ✉ 

Yale University, New Haven, CT, USA

Yueyang Feng ✉ 

Yale University, New Haven, CT, USA

Zhong Shao ✉ 

Yale University, New Haven, CT, USA

Abstract

We present Linearizability Hoare Logic (LHL), the first *mechanized*, *sound*, and *complete* program logic for atomic, set, and interval linearizability. We achieve this by showing soundness and completeness of LHL w.r.t. a more general criterion, *compositional linearizability*, which subsumes all three criteria. We showcase the expressivity of LHL by verifying an exchanger with a set linearizable specification, the elimination-backoff stack built above the exchanger, a lock with an atomic linearized specification, and a write-snapshot object with an interval linearizable specification.

Together with LHL we formalize a modular verification framework for concurrent components based on the theory of compositional linearizability. This allows us to specify components at a high level of abstraction and granularity, and then assemble them into large systems that are correct by construction. As a showcase, we verify the elimination-backoff stack modularly by verifying each of its sub-components against their linearized specifications and then linking them together.

2012 ACM Subject Classification Theory of computation → Program verification; Computing methodologies → Concurrent computing methodologies; Theory of computation → Programming logic

Keywords and phrases Program Logic, Rely-Guarantee, Linearizability, Compositional Verification, Concurrency

Digital Object Identifier 10.4230/LIPIcs.ECOOP.2026.11

Related Version *Extended Version*: <https://flint.cs.yale.edu/publications/lhl.html> [14]

Supplementary Material

Software: <https://github.com/ehatti/LHL/tree/ecoop-camera-ready> [13]

archived at `swb:1:dir:ee421dd7fa7166ad053af08aa59a2cf82c194b33`

Software (ECOOP 2026 Artifact Evaluation approved artifact):

<https://doi.org/10.4230/DARTS.12.1.5>

Funding This work is supported in part by NSF grants 2019285, 2313433, 2442888, and by the Defense Advanced Research Projects Agency (DARPA) under Agreement No. HR00112590130. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

Acknowledgements We would like to thank the anonymous reviewers and artifact evaluation committee for their thoughtful feedback and time spent reviewing our work.



© Eashan Hatti, Arthur Oliveira Vale, Zhongye Wang, Yueyang Feng, and Zhong Shao;
licensed under Creative Commons License CC-BY 4.0

40th European Conference on Object-Oriented Programming (ECOOP 2026).

Editors: Robert Krebbers and Alexandra Silva; Article No. 11; pp. 11:1–11:28



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



1 Introduction

1.1 Compositionality

When designing large computer systems, it is good practice to encapsulate independent functionalities in separate components. This benefits verification, as a well-organized system decomposes into small open components that are completely characterized by their own implementation and their interfaces with the rest of the system.

In this context, given a module M that implements some object with interface F by importing a library with interface E , we call the implemented object the *overlay* of the module and the imported library its *underlay*. To verify the module against its overlay specification, we only consider the relevant underlay specifications, allowing provers to write independent proofs for separate modules. This means that the internals of the underlay module do not matter for other parts of the system: the module is characterized by its underlay and overlay specifications. This mode of verification is only possible in a *compositional* verification framework, where each module can be given a correctness property independent of the rest of the system, and then can be used as a building block to construct a larger system, with no side-conditions for the correctness of their compositions.

In the context of concurrent objects, linearizability [16] has been the gold standard since the 90s. It allows programmers to abstract the behavior of a concurrent object E 's code M_E into a simpler specification V_E that is atomic¹ and easier to reason about. While there are many efforts to enable the mechanized verification of linearizability, few are truly scalable: either there are concurrent objects like the exchanger with non-atomic specifications that cannot be expressed in these pre-existing frameworks due to their over-reliance on atomicity, or the specification cannot be used as a black box encapsulating its code to modularly verify another object that uses this underlay object.

1.2 Compositional Linearizability

Over time, linearizability, which was an inherently atomic correctness criterion, has been extended to handle non-atomic objects [25, 5]. Recently, [27, 26] have also presented an analysis of compositional models of computation which reveals that linearizability, even when fully generalized to handle non-atomic blocking concurrent objects, has an underlying compositional structure, which can be exploited to develop a compositional refinement calculus for the verification of concurrent modules. As an application of their theory, [27] develop an axiomatic technique for showing linearizability of individual traces, and then package it as a rely-guarantee program logic. They call their linearizability criterion and underlying framework *compositional linearizability*.

While an important first step towards the compositional verification of non-atomic concurrent objects, and a convincing showcase of the compositional linearizability theory, the verification technique proposed by [27] could still be improved for the sake of practicality in large systems verification. Some immediate issues are that their framework has not been validated by a mechanization, and while [27] give a paper proof of soundness, they only conjecture its completeness. In this paper, we set out to mechanize the verification technique and formalize the soundness proof of [27], while also proving its completeness. Completeness is the main contribution of our work, and results in the first complete program logic for any of the standard generalizations of linearizability beyond atomicity.

¹ By *atomic* we mean that when a thread makes an invocation it receives its response immediately after, with no interference from other threads.

In doing so, we identified an additional issue for the sake of practicality. To exploit the trace-based theory they had developed earlier, [27] use traces as a universal notion of state. Unfortunately, in practice, using traces as state leads to unnecessary complexity in defining predicates, as simple properties that would be readily available in a concrete representation of state become properties that must be stated inductively over the underlying trace.

So, in order to enable us to verify complex examples, we found it paramount to switch to a state-based representation of proof configurations. It turns out that switching from trace-based reasoning to state-based reasoning is not a simple task. There are two aspects to this. First, [27] use a technique inspired by [16] which they call *possibility reasoning*, where one attempts to maintain a set of *possible* linearizations for the current computation through a representation of linearizations in a structure they call a *possibility*. Unfortunately, some of the possibility update rules of [27] do not readily generalize to a state-based formulation, as they involve manipulating the ordering of events on a trace.

Another aspect of switching to state-based reasoning is that many of the algebraic properties [27] rely on to show the compositionality of the verification framework no longer hold strictly. Technically, their trace-based representation allows them to work up to equality to show several categorical properties of their refinement calculus. We had to generalize these categorical properties to hold only up to notions of equivalence in order to generalize their results to concrete states. In the process, we end up mechanizing a large portion of a generalization of the theory in [27].

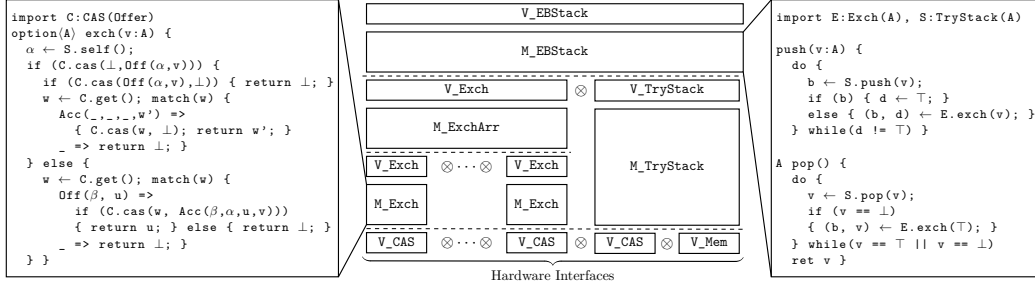
Summary and Main Contributions

- We present the first *mechanized* program logic, Linearizability Hoare Logic (LHL), which supports and unifies the verification of Herlihy-Wing linearizability and other non-atomic linearizabilities into a generalized criterion called *compositional linearizability* [27].
- To mechanize compositional linearizability, we first formalize a compositional model of computation in § 3, formalizing modules, specifications, objects, refinement and composition operations, and proving their algebraic and compositional properties.
- In § 4, we generalize compositional linearizability, originally based solely on traces, to our LTS setting, along with its compositional properties.
- Finally, we present proof rules of our LHL in § 5, which uses a technique called *possibility reasoning* to construct linearizability proofs. We prove its soundness and completeness w.r.t. compositional linearizability.
- Using LHL, we prove several representative examples, including: a modular proof of the elimination-backoff stack [15] in the structure proposed in Fig. 1, to be explained throughout the paper; we mechanize a lock and then a coarse-grained locked racy object (the motivating example used by [27]); to showcase full-blown non-atomicity, we also mechanize an interval-linearizable [5] write-snapshot object.

All results of the paper are mechanized in Rocq and are found as supplemental material. In addition, an extended version [14] of this paper is available containing further details on the elimination-backoff stack and a detailed comparison between our paper and [27].

2 Motivating Example

Throughout the paper, we use the *elimination-backoff stack* [15] (henceforth “EBStack”) to demonstrate our approach.



■ **Figure 1** The exchanger implementation (left), and the hierarchy (middle) of underlay object implementations (M_xxx) and their interfaces (V_xxx) used in the elimination-backoff stack implementation (right): A is the type of values stored in the stack; **Offer** is a data type defined by $\text{Offer} := \text{Off}(\alpha, v) \mid \text{Acc}(\alpha, \beta, v, w)$, where $\alpha, \beta \in \mathcal{T}$ are thread identifiers and $v, w \in A$; The $\perp$ value indicates the failure of the operation in both the exchanger and the EBStack object.

The EBStack can be decomposed into multiple concurrent objects as shown in Fig. 1. Each object implements a specific overlay specification defined as a labelled transition system (LTS), and only uses operations provided in its underlay interfaces. For example, the top-level EBStack object implements the usual stack specification $V_EBStack$ above the exchanger object and a try stack (a stack whose operations may fail). As shown in the code in the right box, the EBStack will first try to perform a push/pop operation through the try stack. If the operation fails, the EBStack tries to eliminate with another operation through the `exch` operation of the exchanger object, which only succeeds if there are concurrent `exch` operations. The EBStack's push/pop operations loop until one of these two steps succeeds. The exchanger implementation is encapsulated in a different module `Exch`, which is built above the CAS memory cells. To increase throughput, an array of exchangers is used instead of the single exchanger in the EBStack [15]. In § 3, we explain in detail the programming language we use to implement these objects and LTS for specifying objects.

The key benefit of this layered approach is that, with proper encapsulation, the implementation and verification of each object become relatively concise, manageable, and, most importantly, independent of other objects. In the hierarchy in Fig. 1, the EBStack code depends only on the V_Exch and $V_TryStack$ interfaces and is decoupled from the implementation and verification of these overlay objects. After verifying the EBStack layer, its proof will remain untouched when the implementation of its underlay objects is modified, as they are all verified to obey the object specifications that are independent of the implementation code. For example, the exchanger array `ExchArr` and the single exchanger `Exch` share the same overlay specification V_Exch because they are meant to implement the same operation with the same functionality, and, as a result, we may use either one as the V_Exch library in the overlay EBStack object without any modification to any implementation code or proof.

This leads to the reason that we are interested in the verification of the elimination-backoff stack. The variations of EBStack shown in Fig. 1 have been verified in other concurrent object verification frameworks [23, 11]. However, none of them achieves the same degree of compositionality shown in this diagram. They all choose to inline the code of the exchanger into the code of the EBStack, which results in complicated EBStack code and proof. Some choose to inline the single exchanger code, which brings another problem: when they want to replace the single exchanger with a more performant implementation like the exchanger array, they need to redo the entire proof of the EBStack. We further discuss the comparison between their proofs and ours in § 6.

The reason for this is that their specifications and correctness criteria cannot support non-atomic objects, such as the exchanger object. An exchange only makes sense if two threads are executing concurrently, which is not possible in an atomic specification. Therefore, the exchanger object does not admit a deterministic atomic linearized specification. Hence, it cannot be verified as a separate module with strict proof encapsulation, and its proof has to be merged into the proof of the overlay stack object. We use the theory of compositional linearizability to specify the exchanger by a set-linearizable LTS, and use a program logic based on possibility reasoning to verify the exchanger and other objects as standalone modules. Compositionality ensures the correctness of the top-level stack w.r.t. its specification $V_EBStack$ in a system running above the hardware interfaces, given the per-module correctness w.r.t. their own overlay specs. We elaborate on this verification technique in § 4 and § 5.

3 Programming Language and Specifications

Our framework verifies concurrent *modules* $M : \text{Mod } E \ F$, which implement operations specified by an *effect signature* F . A module M implements each operation $f \in F$ as a program $M^f : \text{Prog } E \ \text{ar}(f)$, which uses operations from an underlay signature E and returns a value in the return type (specified by $\text{ar}(f)$) of the operation f that it implements. An object over an effect signature E is specified by a state-transition system $V_E : \text{Spec } E$. The core verification task is to show that given a specification $V_E : \text{Spec } E$, called the *underlay* of M , the module $M : \text{Mod } E \ F$ correctly implements its overlay specification $V_F : \text{Spec } F$. We now discuss these notions formally.

3.1 Effect Signatures

An effect signature consists of a set E of effects e and two assignments $\text{par}_E(-) : E \rightarrow \text{Set}$ and $\text{ar}_E(-) : E \rightarrow \text{Set}$, specifying respectively a set of *parameters* and a set of *return* values for each effect $e \in E$. We find it convenient to encapsulate this information in the following notation, whose usage will become clear shortly:

$$E := \{e : \text{par}_E(e) \rightarrow \text{ar}_E(e) \mid e \in E\}$$

When it causes no confusion, we omit the subscript on $\text{par}_E(-)$ and $\text{ar}_E(-)$.

In the context of our paper, effects $e \in E$ are method names, which take as arguments elements of the parameter set $\text{par}(e)$ and return some value in its arity $\text{ar}(e)$. For instance, the signature for the `EBStack` data structure is the following, where A is the type of stack elements and $\mathbf{1}$ is the unit type.

$$\text{EBStack} := \{\text{push} : A \rightarrow \mathbf{1}, \text{pop} : \mathbf{1} \rightarrow A + \mathbf{1}\}$$

Signatures also support a *union* operation, which collects the operations of both signatures into a single, combined signature. Formally:

$$E + F := \{\text{inl } e : \text{par}_E(e) \rightarrow \text{ar}_E(e) \mid e \in E\} \uplus \{\text{inr } e : \text{par}_F(e) \rightarrow \text{ar}_F(e) \mid e \in F\}$$

3.2 Programs

Programs play the role of method bodies within modules in our mechanization. We formalize programs as *interaction trees* [34], which we type as Prog . Interaction trees are coinductively defined data structures with the following syntax:

$$p \in \text{Prog } E \ R := \text{Vis } f(a) \ k \mid \text{Ret } r \mid \text{Tau } p \text{ where } f \in E \quad a \in \text{par}(f) \quad k : \text{ar}(f) \rightarrow \text{Prog } E \ R \quad r \in R$$

A program $p \in \text{Prog } E \ R$ represents a (possibly infinite) series of operations that result in a return value of type R (if finite). The three constructors function as follows.

Vis $f(a) \ k$: Issues an invocation to the operation $f(a)$ of the underlay signature E , and proceeds with the continuation $k(v)$ after receiving a return value $v \in \text{ar}(f)$.

Ret r : Once a program executes all of its operations, it terminates by returning a value r of type R .

Tau p : Represents a silent step and is necessary for looping constructs to satisfy Rocq's productivity checker. We refer readers to [34] for more details.

$\text{Prog } E \ _$ defines a monad for any E , which allows for programs to be written in a monadic style. The program bodies in Fig. 1 are the sugared version of a monadic expression over $\text{Prog } E \ _$. The additional control structures, such as the **if** conditional and the **while** loop constructs used in Fig. 1, may be defined in a standard way.

3.3 Modules

A module $M : \text{Mod } E \ F$ implements operations of a signature F , using the operations of the signature E . A module $M : \text{Mod } E \ F$ consists of a mapping from operations $f \in F$ and arguments $a \in \text{par}(f)$ to programs $M^{f(a)} : \text{Prog } E \ \text{ar}(f)$, implementing a call $f(a)$ as a program operating over E and returning a value in its return type $\text{ar}(f)$. In other words, a module $M : \text{Mod } E \ F$ is a collection $M : \prod_{f \in F} (\text{par}(f) \rightarrow \text{Prog } E \ \text{ar}(f))$. For instance, $M_{\text{EBS\textsubscript{tack}}}$ is the mapping taking $\text{push}(v)$ to the program described in Fig. 1 and similarly for $\text{pop}()$. Nonetheless, we write $M^{f(a)} : \text{Prog } E \ \text{ar}(f)$ for the program corresponding to $f \in F$ and $a \in \text{par}(f)$ in M , and $M^f : \text{par}(f) \rightarrow \text{Prog } E \ \text{ar}(f)$.

Modules may be composed vertically by interaction tree substitution. Given modules $M : \text{Mod } E \ F$ and $N : \text{Mod } F \ G$, their vertical composition $M :> N : \text{Mod } E \ G$ implements the operation $g(a)$ of G by running the corresponding program $N^{g(a)}$ with method invocations $f(a')$ to F replaced by their corresponding programs $M^{f(a')}$. We refer readers to the mechanization for the exact definition. Crucially, the vertical composition operation admits an identity element:

$$\text{idM}_E : \text{Mod } E \ E := (\text{idProg}_e)_{e \in E} \quad \text{where } \text{idProg}^e := \lambda(a : \text{par}(e)). \text{Vis } e(a) \ \text{Ret}$$

The vertical composition operation $(: >)$ and its identity element satisfy the following properties:

$$(1) \text{idM} :> M \approx M \quad (2) M :> \text{idM} \approx M \quad (3) (M_1 :> M_2) :> M_3 \approx M_1 :> (M_2 :> M_3)$$

where $\approx$ is an equivalence relation defined on modules as the pointwise equivalence of programs up to removal of silent steps (**Tau**). The first two properties show that idM is the neutral element for vertical composition, up to $\approx$ equivalence, and the third property is its associativity.

3.4 Specifications

We now give a formal definition of *labelled transition system* (LTS), the *specifications* we use for objects. Recall that a module $M : \text{Mod } E \ F$ depends on the operations provided by E to implement the operations of F . To define the operational semantics of a module, we first define the behavior of the operations in E using an LTS.

First, we must define what kinds of events we are interested in. We assume as given a set of thread names $\mathcal{T}$, which is a parameter over the whole model. We use a set of the form $\{i \in \mathbb{N} \mid i < T\}$, where $T \in \mathbb{N}$ is the total number of threads, as $\mathcal{T}$. In a multicore setting, $\mathcal{T}$ is the set of cores available; in a multi-threaded setting, it is the set of available threads. While users can fix the whole set $\mathcal{T}$ when verifying an object, we choose to parameterize our theories and proofs over $\mathcal{T}$. As a result, our proofs of linearizability universally quantify over such finite sets of thread names.

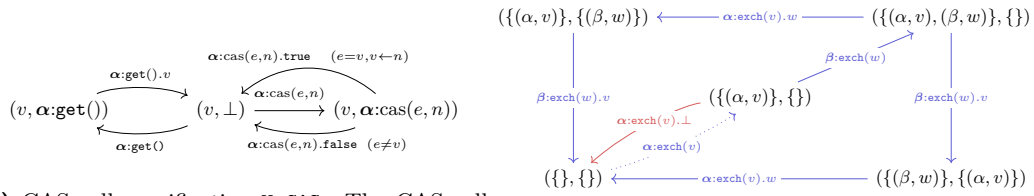
Then, an event of the signature $\mathbf{E}$ consists of

$$\mathbf{TEvent} \mathbf{E} := \alpha:f(a) \mid \alpha:f(a).v$$

where $\alpha \in \mathcal{T}$, $f \in \mathbf{E}$, $a \in \mathbf{par}(f)$ and $v \in \mathbf{ar}(f)$. An event is either $\alpha:f(a)$, a call to an operation of $\mathbf{E}$, or $\alpha:f(a).v$, a return event tagged with the thread that issued it. Then, a labelled transition system $\mathbf{V}$ over $\mathbf{E}$, written $\mathbf{V} : \mathbf{Spec} \mathbf{E}$, is defined as a triple $(S, \rightarrow_{\mathbf{V}}, s_0)$ of a set of states S , a transition relation $\rightarrow_{\mathbf{V}} \subseteq S \times \mathbf{TEvent} \mathbf{E} \times S$, and an initial state $s_0 \in S$.

Unlike many approaches for modeling concurrent specifications, which assume operations are logically atomic, the call and return events are separate transitions in our specifications. This is necessary to model objects that are not linearizable to atomic specifications. We do, however, require that the labelled transition system be *locally sequential*, i.e., the projection to each thread is sequential (it strictly alternates between invocations and responses). As a result, we must carry some extra state, like the operation that is currently executing (e.g., Fig. 2a), so we prevent other calls from happening until the current pending call is completed.

We now give a couple of examples of linearized state transition systems used in the verification of the EBStack. In our state diagrams, such as Fig. 2a, we use conditionals and state updates in the labels, and depict the transitions for abstracted thread names α, β , to mean that for any thread α a certain transition exists (if threads have different names they are distinct $\alpha \neq \beta$). This allows us to make the diagrams finite so they may be depicted succinctly on paper.



(a) CAS cell specification $\mathbf{V_CAS}$. The CAS cell has two operations: $\text{get}()$ for retrieving the stored value v , and $\text{cas}(e, n)$ for performing the compare-and-swap operation.

(b) Exchanger specification $\mathbf{V_Exch}$. The exchanger operation exch allows two threads to attempt to exchange values with other threads.

Figure 2 CAS and Exchanger Linearized Specifications.

► **Example 1** (Atomic CAS Cell LTS). We begin with an atomic specification encoded by the LTS in Fig. 2a for a CAS cell object. Its states are tuples consisting of the stored value and the ongoing pending call. A non-empty pending call component restricts the next transition to be labelled with the matching return event. The two transitions on the left ensure that a get operation $\alpha:\text{get}()$ always gets the stored value v in the return event $\alpha:\text{get}().v$. The three transitions on the right specify that a CAS operation $\alpha:\text{cas}(e, n)$ receives **true** as a response only when the stored value is e and stores the new value n to it, and otherwise, it receives **false** and does not modify the stored value. This LTS specifies an atomic object because, after an invocation by thread α , the next transition must be a response by the same thread α . Invocation transitions are possible only when the ongoing operation is $\perp$, resulting in only sequential executions.

► **Example 2** (Set-Linearizable Exchanger LTS). Fig. 2b defines the LTS specification for the exchanger object. Its LTS state consists of two sets over tuples of a thread name and an exchanged value. The left and right sets are the “offer set” and “accept set”, containing respectively the exchange offers that threads have proposed and accepted. In a successful exchange, two threads will add their offers to the offer set, after which one thread accepts the other offer. The remaining thread will then accept another offer and empty the set. Tracing through the *blue* arrows, we can extract the execution of this successful exchange as trace (1).

$$\alpha:\text{exch}(v) \cdot \beta:\text{exch}(w) \cdot \alpha:\text{exch}(v).w \cdot \beta:\text{exch}(w).v \quad (1)$$

In a failed exchange, the offering thread will revoke its offer and return $\perp$. This will step through the dotted blue arrow and the *red* arrow as trace (2).

$$\alpha:\text{exch}(v) \cdot \alpha:\text{exch}(v).\perp \quad (2)$$

All possible executions of the exchanger in the specification are repetitions of trace (1) and trace (2). In trace (1), calls are not necessarily followed immediately by their returns – the calls and returns of both threads α and β are interleaved with each other. As a result, this specification is **not** atomic linearizable but is *set linearizable* [25]: calls and returns in the exchanger’s execution traces always come as a set of at most two calls followed by a set of matching returns.

3.5 Traces and Refinement

We define a trace over a signature $\mathbf{E}$ as a list $t \in (\mathbf{TEvent} \ \mathbf{E})^*$ of thread events. We write $\mathbf{Trace} \ \mathbf{E}$ for the set of traces over $\mathbf{E}$. The LTS specification $\mathbf{V} : \mathbf{Spec} \ \mathbf{E}$ specifies the behavior of an object through $\llbracket \mathbf{V} \rrbracket$, a set of traces defined as follows: we first define a path in $\mathbf{V} = (S, \rightarrow_{\mathbf{V}}, s_0)$ and then take all paths starting from the initial state as the set of behaviors.

$$\frac{s \in S}{s \xrightarrow{\epsilon}_{\mathbf{V}} s} \quad \frac{s_1 \xrightarrow{\text{ev}}_{\mathbf{V}} s_2}{s_1 \xrightarrow{\text{ev}}_{\mathbf{V}} s_2} \quad \frac{s_1 \xrightarrow{t}_{\mathbf{V}} s_2 \quad s_2 \xrightarrow{t'}_{\mathbf{V}} s_3}{s_1 \xrightarrow{t \cdot t'}_{\mathbf{V}} s_3} \quad \llbracket \mathbf{V} \rrbracket := \{t \in (\mathbf{TEvent} \ \mathbf{E})^* \mid \exists s \in S. s_0 \xrightarrow{t}_{\mathbf{V}} s\}$$

Our notion of refinement is then just behavior containment:

$$\mathbf{V} \sqsupseteq \mathbf{V}' \iff \llbracket \mathbf{V} \rrbracket \subseteq \llbracket \mathbf{V}' \rrbracket \quad \mathbf{V} \equiv \mathbf{V}' \iff \mathbf{V} \sqsupseteq \mathbf{V}' \wedge \mathbf{V}' \sqsupseteq \mathbf{V}$$

It is straightforward to check that refinement is reflexive and transitive, making it a preorder.

3.6 Operational Semantics

Given an underlay specification $\mathbf{V_E} : \mathbf{Spec} \ \mathbf{E}$, where $\mathbf{V_E} = (S, \rightarrow_{\mathbf{V}}, s_0)$, and a module $M : \mathbf{Mod} \ \mathbf{E} \ \mathbf{F}$, we define the behaviors of M running on top of $\mathbf{V_E}$ operationally as a concrete LTS ($\mathbf{V_E} \triangleright M$). In § 4, we define the linearizable relation of the concrete LTS w.r.t. the specification LTS (those in Fig. 2).

The operational semantics models executions under an arbitrary scheduler and client to the overlay object $\mathbf{F}$. An arbitrary client may have a thread call any operation available to it, with any argument. Hence, each operational semantics rule non-deterministically chooses a thread to take a step. This ensures we verify our modules as open components.

The states $(ths, s) \in \mathbf{InterState}$ (interaction state) of the operational semantics are pairs of a *thread state* $ths \in \mathcal{T} \rightarrow \mathbf{TState} \ \mathbf{E} \ \mathbf{F}$, which is a map from thread names to the continuation of that thread, and an *underlay LTS state* $s \in S$. The continuation of each thread is defined as

$$\mathbf{TState} \ \mathbf{E} \ \mathbf{F} := \mathbf{Idle} \mid \mathbf{Cont} \ f(a) \ p \mid \mathbf{UCall} \ f(a) \ g(b) \ k$$

$$\begin{array}{c}
\text{OCALL} \quad \frac{ths(\alpha) = \text{Idle}}{(ths, s) \xrightarrow{\alpha:f(a)} (ths[\alpha \mapsto \text{Cont } f(a) M^{f(a)}], s)} \\
\text{UCALL} \quad \frac{ths(\alpha) = \text{Cont } f(a) (\text{Vis } g(b) k) \quad s \xrightarrow{\alpha:g(b)}_{V_E} s'}{(ths, s) \xrightarrow{\alpha:g(b)} (ths[\alpha \mapsto \text{UCall } f(a) g(b) k], s')} \\
\text{URET} \quad \frac{ths(\alpha) = \text{UCall } f(a) g(b) k \quad s \xrightarrow{\alpha:g(b).v}_{V_E} s'}{(ths, s) \xrightarrow{\alpha:g(b).v} (ths[\alpha \mapsto \text{Cont } f(a) k(v)], s')} \\
\text{SILENT} \quad \frac{ths(\alpha) = \text{Cont } f(a) (\text{Tau } p)}{(ths, s) \xrightarrow{\epsilon} (ths[\alpha \mapsto \text{Cont } f(a) p], s)}
\end{array}$$

■ **Figure 3** Operational Semantics.

where $f \in F$, $a \in \text{par}(f)$, $p \in \text{Prog } E \text{ ar}(f)$, $g \in E$, $b \in \text{par}(g)$, $k : \text{ar}(g) \rightarrow \text{Prog } E \text{ ar}(f)$. We use $ths[\alpha \mapsto p]$ as the notation for updating the value of α to p in the map ths .

Fig. 3 defines the operational semantics rules. OCALL selects an idle thread, non-deterministically chooses an operation $f \in F$ and an argument $a \in \text{par}(f)$, and changes the continuation for α to the code for $f(a)$ defined as $M^{f(a)}$. UCALL locates a thread α whose next step in the continuation is an underlay call $g(b)$, performs the call step $\alpha:g(b)$ in the underlay object V_E , and updates the underlay LTS state to s' accordingly. URET is analogous, except that it takes a return transition and continues with the continuation as specified by k . ORET matches on a thread whose continuation indicates it is ready to return from the overlay operation, and performs the overlay return by updating the continuation to Idle . SILENT simply skips a Tau with no change to the state.

We then define the concrete LTS of running M on top of V_E as $V_E \triangleright M : \text{Spec } F := (\text{InterState } F \ V_E, \rightarrow_{V_E \triangleright M}, (ths_0, s_0))$, where ths_0 is the constant mapping to Idle and $m \in \text{TEvent } F$, and transitions are given by the following relation

$$\begin{aligned}
(ths, s) \xrightarrow{m}_{V_E \triangleright M} (ths', s') &\iff \exists ths'' \in \mathcal{T} \rightarrow \text{TState } E \ F, s'' \in S, t \in (\text{TEvent } E)^* \\
&\quad (ths, s) \xrightarrow{m} (ths'', s'') \wedge (ths'', s'') \xrightarrow{t} (ths', s')
\end{aligned}$$

Notice that only the overlay events $m \in \text{TEvent } F$ are visible in $V_E \triangleright M$. The definition allows for a visible step m followed by several steps in the underlay object V_E , as enforced by the requirement that $t \in (\text{TEvent } E)^*$. The underlay event trace is hidden behind an existential quantifier, making these implicit silent steps.

The following are some important properties of the operation $\triangleright$. The first two properties are monotonicity of vertical composition w.r.t. $\approx$ and $\sqsubseteq$, while the third property encodes both associativity of $\triangleright$ and $:\triangleright$, as well as compatibility of the two operations.

$$\begin{aligned}
(1) \ M \approx M' &\Rightarrow V \triangleright M \equiv V \triangleright M' & (2) \ V' \sqsubseteq V &\Rightarrow V' \triangleright M \sqsubseteq V \triangleright M \\
(3) \ V \triangleright (M : \triangleright N) &\equiv (V \triangleright M) \triangleright N
\end{aligned}$$

3.7 Horizontal Composition

Both modules and specifications can be horizontally composed. For instance, given modules $M_L : \text{Mod } E_L \ F_L$ and $M_R : \text{Mod } E_R \ F_R$ we can define their horizontal composition $M_L + M_R : \text{Mod } (E_L + E_R) \ (F_L + F_R)$, similarly to how we did with effect signatures, as the union of the two collections:

$$(M_L + M_R)^{f'} := \begin{cases} \text{mapProg}(\lambda x. \text{inl } x) M_L^f & f' = \text{inl } f \\ \text{mapProg}(\lambda x. \text{inr } x) M_R^f & f' = \text{inr } f \end{cases}$$

11:10 A Complete Program Logic for Compositional Linearizability

where $\text{mapProg } (\lambda x. \text{inl } x)$ replaces every event $\text{Vis } e$ in M_L^f by $\text{Vis } (\text{inl } e)$ (and similarly for the other branch).

For specifications $V_L : \text{Spec } E_L$ and $V_R : \text{Spec } E_R$ we can define a specification $V_L \otimes V_R$ given by the (locally sequential) asynchronous product of the two transition systems. We refer the reader to the mechanization for the definition. Below, we show the key compositional properties of $- \otimes -$:

$$\frac{V_L \sqsupseteq V_L' \quad V_R \sqsupseteq V_R'}{V_L \otimes V_R \sqsupseteq V_L' \otimes V_R'} \quad \frac{V_L \otimes V_R \sqsupseteq V_L' \otimes V_R'}{V_L \sqsupseteq V_L'} \quad \frac{V_L \otimes V_R \sqsupseteq V_L' \otimes V_R'}{V_R \sqsupseteq V_R'}$$

$$\text{id}_{M_E} + \text{id}_{M_F} = \text{id}_{M_{E+F}}$$

$$(V_L \triangleright M_L) \otimes (V_R \triangleright M_R) \equiv (V_L \otimes V_R) \triangleright (M_L + M_R)$$

The first three properties show that horizontal composition defines an order-isomorphism for specification refinement. The last two properties show that the operation is functorial.

4 Linearizability and Possibilities

We now define the notion of linearizability used in the paper (§ 4.2), which is based on the compositional linearizability formulation by [27]. Then, we go over how we implement one of the main techniques for proving linearizability: possibility reasoning (§ 4.4).

4.1 Background

Safety for sequential objects is usually given by functional correctness. For instance, a counter object provides operations $\text{get} : 1 \rightarrow \mathbb{N}$, taking unit as argument and returning some natural number, and $\text{inc} : 1 \rightarrow 1$, taking unit as argument and returning unit. An implementation of a counter is correct as long as get always returns the number of previous inc operations. This means that given the trace:

$$s = \text{inc} \cdot \text{ok} \cdot \text{get}$$

of a correct counter implementation, we are guaranteed that the next event is a return of 1 to the operation get . However, for a concurrent counter implementation, given the trace:

$$\alpha_1:\text{inc} \cdot \alpha_1:\text{ok} \cdot \alpha_1:\text{get}$$

any number of increment operations by other threads might take effect:

$$\alpha_1:\text{inc} \cdot \alpha_1:\text{ok} \cdot \alpha_1:\text{get} \cdot \alpha_2:\text{inc} \cdot \alpha_2:\text{ok} \dots \alpha_k:\text{inc} \cdot \alpha_k:\text{ok} \cdot \alpha_1:n$$

so that n might be any number between 1 and k . What makes a concurrent counter object correct is therefore a more challenging question, and expressing “functional” correctness for concurrent objects remained an open question until the 90s, when linearizability [16] was proposed.

Linearizability asks that every trace generated by the concurrent counter implementation, such as trace s above, be *linearizable* with respect to some trace of a correct sequential counter. This is formalized by defining a partial order called *happens-before*, which associates to a trace s the partial order $\text{hb}(s)$ defined as the smallest partial order such that $e \prec_{\text{hb}(s)} e'$ whenever e and e' are events by the same thread and e appears before e' in s , or when e is a return event and e' is an invocation event. A trace s is then said to be linearizable w.r.t. an

atomic trace t when there is a way to complete some pending invocations of s (by adding extra return events), then remove all remaining pending invocations, so as to obtain a trace s' satisfying that $\text{hb}(s')$ is refined by $\text{hb}(t)$ (in the sense that if $e \prec_{\text{hb}(s')} e'$ then $e \prec_{\text{hb}(t)} e'$). Atomic means that every invocation is immediately followed by its return, or, equivalently, that t is locally sequential (every thread alternates invocations and returns) and $\prec_{\text{hb}(t)}$ is a linear order. An implementation is then said to be linearizable w.r.t. a specification if all traces generated by the implementation are linearizable w.r.t. the specification.

The requirement that t , and consequently specifications, be atomic ensures that concurrent data structures such as stacks, queues, and counters are all specified by essentially the same specification as the corresponding sequential data structures. This requirement, however, proves too restrictive, as many reasonable concurrent objects admit no corresponding sequential object. For example, the exchanger in § 2 is a standard synchronization primitive (for instance, available as part of Java's standard libraries [4]). Nonetheless, it admits no reasonable atomic specification, intuitively because an exchange can only happen if two threads are running concurrently. To remedy this, *set linearizability* was proposed [25]. Set linearizability only modifies the definition of linearizability by requiring that the traces of the specification be of the form:

$$I_1 \cdot R_1 \cdot I_2 \cdot R_2 \cdot \dots \cdot I_n \cdot R_n$$

where each I_k is a sequence of invocations and each R_k is a sequence of responses containing *all* and *only* the responses to the invocations in I_k (except for R_n , which may contain only *some* such responses). In addition, one requires that if

$$I_1 \cdot R_1 \cdot I_2 \cdot R_2 \cdot \dots \cdot I_n \cdot R_n$$

is in the specification, then so is

$$I'_1 \cdot R'_1 \cdot I'_2 \cdot R'_2 \cdot \dots \cdot I'_n \cdot R'_n$$

where each I'_k (resp. R'_k) is some reordering of I_k (resp. R_k). This means that we may see these traces as sequences of sets of invocations, each followed by the set of returns to that set of invocations.

This modification to linearizability allows it to specify many concurrent objects whose behaviors involve some kind of synchronization, such as exchangers, barriers, and blocking stacks (stacks where `pop` blocks on the empty stack). We will see that it also helps in specifying objects beyond synchronization primitives, such as the `TryStack`. However, there are concurrent objects and certain distributed tasks whose consistency is even weaker than that provided by set linearizability. For example, it is possible to implement a write-snapshot object [3]² validating the trace:

$$s' = \alpha:\text{ws}(3) \cdot \beta:\text{ws}(5) \cdot \alpha:\text{ws}(3).\{3, 5\} \cdot \gamma:\text{ws}(2) \cdot \beta:\text{ws}(5).\{3, 5, 2\}$$

Here, when α takes its snapshot, β 's value of 5 is already visible, but by the time β manages to take its own snapshot, γ 's value is now visible too. Because of this, the return to β 's call cannot happen together with α 's return, and hence set linearizability cannot capture this object's consistency.

² A write-snapshot provides a single operation `ws` that takes a value as argument and returns a set of written values as output.

11:12 A Complete Program Logic for Compositional Linearizability

To accommodate such objects, interval linearizability [5] was proposed, which further relaxes the notion of specification in linearizability's definition so that in a trace

$$I_1 \cdot R_1 \cdot I_2 \cdot R_2 \cdot \dots \cdot I_n \cdot R_n$$

R_k is now allowed to contain returns to pending invocations from I_l for any $l \leq k$, and in particular it may contain only *some* of the returns in any I_l (including I_k). The requirement that reorderings within I_k or R_k are all valid remains. This has been shown to be a *complete* correctness criterion for distributed tasks [5, 12]. Interval linearized specifications therefore allow traces with a significant degree of concurrency, such as s' above. Nonetheless, these specifications may still be simpler than the traces generated by the implementation. The write-snapshot we discuss and verify in our mechanization admits an interval linearized specification that is deterministic in the sense that given a trace:

$$I_1 \cdot R_1 \cdot \dots \cdot I_k$$

then in any two extensions

$$I_1 \cdot R_1 \cdot \dots \cdot I_k \cdot R_k \quad \text{and} \quad I_1 \cdot R_1 \cdot \dots \cdot I_k \cdot R'_k$$

where R_k and R'_k contain only responses, if $\alpha:\mathbf{ws}(v).S$ appears in R_k then either $\alpha:\mathbf{ws}(v).S$ also appears in R'_k or R'_k can be extended with $\alpha:\mathbf{ws}(v).S$. In other words, any returns that can happen for any given thread are unique up to the ordering of the responses. This makes the linearized specification significantly easier to work with for a client of the write-snapshot object than using all the traces generated by the implementation, which present more concurrency and non-determinism.

Nonetheless, interval linearizability still does not accommodate all concurrent objects. This last issue is more subtle, and we recommend that the interested reader consult [27] for a more thorough treatment. Succinctly, the requirement that *all* pending invocations that have not been completed must be removed in the linearization is too strong for certain concurrent objects. There are circumstances where an invocation is *detectable* to other threads even though its response is still not enabled. Compositional linearizability [26, 27] further weakens interval linearizability by allowing some invocations to be completed, some to be removed, and some to remain in the linearization. This makes compositional linearizability complete for concurrent objects. However, perhaps the main advance of compositional linearizability is that it gives a novel treatment of the theory of linearizability showing that, within a compositional model of computation, linearizability can be characterized entirely in terms of composition and refinement. In this paper, we adopt compositional linearizability as our notion of correctness and mechanize many aspects of the comprehensive theory of [27]. In particular, they present rather algebraic proofs of all the key properties of linearizability based on a different (but equivalent) formulation that circumvents the use of partial orders or rewrite systems. We have found that this reformulation makes mechanization tractable despite the significantly more challenging setting of working with non-atomic and potentially blocking specifications. At the same time, we find it productive to formulate their ideas using LTSs rather than strategies (final coalgebras for these LTSs). To do so, we made several improvements to their proof technique based on possibilities. The remainder of this section concerns our mechanization of compositional linearizability and possibilities, ultimately leading to our proof technique in § 5.

4.2 Linearizability

The observation at the core of [27] is that, in the context of a compositional model for computation, the linearizability theory can be entirely derived from the definition of the identity for composition of modules. They refer to their notion of linearizability as *compositional linearizability* and demonstrate that it is a conservative generalization of Herlihy-Wing linearizability [16], set linearizability [25], and interval linearizability [5], as it becomes equivalent to these criteria under appropriate restrictions.

In our mechanization, we choose to formalize compositional linearizability using its refinement-based definition, which can be defined using idM as follows:

► **Definition 3.** *Given a specification $V : \text{Spec } E$, we define $K V : \text{Spec } E$ by $K V := V \triangleright \text{idM}$.*

We say a concrete LTS $V' : \text{Spec } E$ is linearizable w.r.t. $V : \text{Spec } E$, written $V' \rightsquigarrow V$ when $V' \sqsubseteq K V$.

Intuitively, the reason why this definition agrees with the usual linearizability definition is that the set of behaviors of $K V$ is exactly the set of all traces linearizable w.r.t. some behavior of V . The operations of E made by the overlay to idM form a bracket around the corresponding underlay operations. Because of this, happens-before ordering is preserved from the overlay of idM to its underlay. The fact that pending calls may be removed or completed in a trace corresponds to the fact that in the operational semantics of $V \triangleright \text{idM}$, one can be in a state where the call has been made in the overlay of idM but not in the underlay yet. There are states where a return has been made to the underlay operation in idM but has not returned to the overlay yet. Therefore, $K V$ defined above contains all traces linearizable w.r.t. the specification. We refer readers to [27] for a detailed treatment of how this definition relates to previous definitions of linearizability.

Our definition of linearizability enjoys observational refinement and locality (Theorem 4). They follow from generalizing the proofs of compositional linearizability from [27] to handle our more heterogeneous compositional structure, discussed in § 3.

► **Theorem 4** (Observational Refinement and Locality). *For any $V_{-E'}, V_{-E} \in \text{Spec } E$, $V_{-F'}, V_{-F} \in \text{Spec } F$, and $M \in \text{Mod } E \ G$, we have observational refinement (1) and locality (2):*

- (1) *Behaviors of running any module M above the concrete LTS $V_{-E'}$ are contained in those of running the same M above the linearized specification V_{-E} , i.e., $V_{-E'} \rightsquigarrow V_{-E} \implies (V_{-E'} \triangleright M) \sqsubseteq (V_{-E} \triangleright M)$.*
- (2) *A system of multiple objects is linearizable iff all of these objects are separately linearizable, i.e., $V_{-E'} \rightsquigarrow V_{-E} \wedge V_{-F'} \rightsquigarrow V_{-F} \iff V_{-E'} \otimes V_{-F'} \rightsquigarrow V_{-E} \otimes V_{-F}$.*

4.3 Compositionality

Through compositional linearizability [27], LHL gains horizontal and vertical composition without any additions to the logic. LHL's soundness provides linearizability proofs of different modules, which can be horizontally and vertically composed together to obtain the linearizability of a larger system. These results are obtained by combining locality and observational refinement with the compositional structure of the model defined in § 3.

► **Theorem 5** (Horizontal Comp.). *If $(V_{-E_1} \triangleright M_{-F_1}) \rightsquigarrow V_{-F_1}$ and $(V_{-E_2} \triangleright M_{-F_2}) \rightsquigarrow V_{-F_2}$, then $((V_{-E_1} \otimes V_{-E_2}) \triangleright (M_{-F_1} \otimes M_{-F_2})) \rightsquigarrow (V_{-F_1} \otimes V_{-F_2})$.*

► **Theorem 6** (Vertical Comp.). *If $(V_{-E} \triangleright M_{-E}) \rightsquigarrow V_{-F}$ and $(V_{-F} \triangleright M_{-F}) \rightsquigarrow V_{-G}$, then $(V_{-E} \triangleright (M_{-E} \triangleright M_{-F})) \rightsquigarrow V_{-G}$.*

Horizontal composition enables importing multiple objects together through the tensor operator ($\otimes$) and using them concurrently. The locality of compositional linearizability guarantees that both objects will preserve their original specifications, and users of both objects can enjoy the interface $F_1 + F_2$ with both sets of operations available.

Vertical composition ($\triangleright$) “links” the overlay object’s code with the imported underlay object’s code. Since vertical composition ($\triangleright$) is compatible with linking ($:\triangleright$), the observational refinement [27] ensures running the linked code $M_E : \triangleright M_F$ above the lower-level specification V_E still produces traces linearizable to the top-level specification V_F .

This is how we “glue” together the individual verified components in Fig. 1 into the whole implementation of the EBStack to obtain the correctness of the entire system.

4.4 Possibility Reasoning

To prove linearizability, we employ *possibility reasoning*. [27] have shown a bisimulation between possibilities and their formulation of the operational semantics of `idM` running on top of V_F , implying that possibility reasoning is sound and complete for compositional linearizability proofs. In our setting, we redesign the possibility reasoning technique to contend with our use of state as explained in § 1.

Assuming we want to show M running above V_E is linearizable w.r.t. V_F , i.e., $(V_E \triangleright M) \sim V_F$, a possibility $\rho \in \text{Poss}$ is defined as a triple $\langle \rho_S, \rho_C, \rho_R \rangle$ of (1) the linearized state ρ_S in V_F , (2) a partial map ρ_C from thread names $\alpha \in \mathcal{T}$ to call events of F , and (3) a partial map ρ_R from thread names $\alpha \in \mathcal{T}$ to return events of F . We write $\rho_C(\alpha)$ for the event that ρ_C maps α to; $\rho_C(\alpha) = \perp$ when it is undefined; and $\rho_C[\alpha \mapsto m]$ for substitution of m for $\rho_C(\alpha)$; and similarly for ρ_R . The mapping ρ_C stores pending calls that have happened in the concrete execution but have not been performed in the linearized state s ; ρ_R stores returns that have been made in the linearized state ρ_S , but not in the concrete execution. Therefore, from the perspective of a particular thread α , a possibility can be in one of four states:

Idle: The last transition by α in V_F was a return, and both $\rho_C(\alpha) = \perp$ and $\rho_R(\alpha) = \perp$.

Pending Call: α ’s last transition in V_F was a return, $\rho_C(\alpha) = m$ for some $m \in F$ and $\rho_R(\alpha) = \perp$.

Call Done: The last transition by α in V_F was a call $\alpha:m$, $\rho_C(\alpha) = m$ and $\rho_R(\alpha) = \perp$.

Pending Return: The last transition by α in V_F was a return $\alpha:m.n$, $\rho_C(\alpha) = m$ and $\rho_R(\alpha) = m.n$.

We formalize possibility updates, the axioms to manipulate possibilities, as a transition system whose states are possibilities over V_F and transitions are given by:

$$\begin{array}{c}
\text{INVOKE} \\
\frac{\rho_C(\alpha) = \perp \quad \rho_R(\alpha) = \perp}{\langle \rho_S, \rho_C, \rho_R \rangle \dashrightarrow \langle \rho_S, \rho_C[\alpha \mapsto m], \rho_R \rangle} \\
\\
\text{CCOMMIT} \\
\frac{\rho_S \xrightarrow{\alpha:m} t \quad \rho_C(\alpha) = m \quad \rho_R(\alpha) = \perp}{\langle \rho_S, \rho_C, \rho_R \rangle \dashrightarrow \langle t, \rho_C, \rho_R \rangle} \\
\\
\text{RETURN} \\
\frac{\rho_C(\alpha) = m \quad \rho_R(\alpha) = n}{\langle \rho_S, \rho_C, \rho_R \rangle \dashrightarrow \langle \rho_S, \rho_C[\alpha \mapsto \perp], \rho_R[\alpha \mapsto n] \rangle} \\
\\
\text{RCOMMIT} \\
\frac{\rho_S \xrightarrow{\alpha:m.n} t \quad \rho_C(\alpha) = m \quad \rho_R(\alpha) = \perp}{\langle \rho_S, \rho_C, \rho_R \rangle \dashrightarrow \langle t, \rho_C, \rho_R[\alpha \mapsto n] \rangle} \\
\\
\bar{\rho} \rightarrow \bar{\sigma} \iff \forall \sigma \in \bar{\sigma}. \exists \rho \in \bar{\rho}. \rho \dashrightarrow^* \sigma
\end{array}$$

■ **Figure 4** Possibility update rules.

	$\langle ([], \perp), \emptyset, \emptyset \rangle$
$\rightarrow$ (INVOKE)	$\langle ([], \perp), [\alpha_1 \mapsto \text{push}(1)], \emptyset \rangle$
$\rightarrow$ (INVOKE)	$\langle ([], \perp), [\alpha_1 \mapsto \text{push}(1)][\alpha_2 \mapsto \text{push}(2)], \emptyset \rangle$
$\rightarrow$ (CCOMMIT)	$\langle ([], \alpha_i:\text{push}(i)), [\alpha_1 \mapsto \text{push}(1)][\alpha_2 \mapsto \text{push}(2)], \emptyset \rangle$
$\rightarrow$ (RCOMMIT)	$\left\langle (i :: [], \perp), \begin{bmatrix} \alpha_1 \mapsto \text{push}(1), \\ \alpha_2 \mapsto \text{push}(2) \end{bmatrix}, [\alpha_i \mapsto \text{push}(i) \cdot ()] \right\rangle$
$\rightarrow$ (CCOMMIT)	$\left\langle \left(\begin{array}{c} i :: [], \\ \alpha_j:\text{push}(j) \end{array} \right), \begin{bmatrix} \alpha_1 \mapsto \text{push}(1), \\ \alpha_2 \mapsto \text{push}(2) \end{bmatrix}, [\alpha_i \mapsto \text{push}(i) \cdot ()] \right\rangle$
$\rightarrow$ (RCOMMIT)	$\left\langle (j :: i :: [], \perp), \begin{bmatrix} \alpha_1 \mapsto \text{push}(1), \\ \alpha_2 \mapsto \text{push}(2) \end{bmatrix}, \begin{bmatrix} \alpha_i \mapsto \text{push}(i) \cdot (), \\ \alpha_j \mapsto \text{push}(j) \cdot () \end{bmatrix} \right\rangle$
$\rightarrow$ (RETURN)	$\langle (j :: i :: [], \perp), [\alpha_j \mapsto \text{push}(j)], [\alpha_j \mapsto \text{push}(j) \cdot ()] \rangle$
$\rightarrow$ (RETURN)	$\langle (j :: i :: [], \perp), \emptyset, \emptyset \rangle$
$\rightarrow$ (INVOKE)	$\langle (j :: i :: [], \perp), [\alpha_3 \mapsto \text{pop}()], \emptyset \rangle$
$\rightarrow$ (CCOMMIT)	$\langle (j :: i :: [], \alpha_3:\text{pop}()), [\alpha_3 \mapsto \text{pop}()], \emptyset \rangle$
$\rightarrow$ (RCOMMIT)	$\langle (i :: [], \perp), [\alpha_3 \mapsto \text{pop}()], [\alpha_3 \mapsto \text{pop}() \cdot j] \rangle$

■ **Figure 5** Example of possibility-based linearizability proof.

Intuitively, a possibility encodes a proof of linearizability. The INVOKE and RETURN model the events of the concrete trace, and CCOMMIT and RCOMMIT the corresponding events in the linearization. The axioms on possibility updates ensure that as long as the initial possibility is valid, the updated possibility always corresponds to some *valid* linearization, i.e., (1) only call events that happened in the concrete execution can be added to the possibility, (2) only transitions respecting local sequentiality can be made to the possibility, and (3) return events missing in the concrete execution can be added to complete pending calls in the possibility. The remaining obligation to show the consistency of linearized return values and actual returned values will be addressed by the program logic. For instance, a possible proof that the trace t' below is linearizable w.r.t. $t_{i,j}$ corresponds to the sequence of possibility updates shown in Fig. 5.

$$t' = \alpha_1:\text{push}(1) \cdot \alpha_2:\text{push}(2) \cdot \alpha_1:\text{push}(1) \cdot () \cdot \alpha_2:\text{push}(2) \cdot () \cdot \alpha_3:\text{pop}()$$

$$t_{i,j} = \alpha_i:\text{push}(i) \cdot \alpha_i:\text{push}(i) \cdot () \cdot \alpha_j:\text{push}(j) \cdot \alpha_j:\text{push}(j) \cdot () \cdot \alpha_3:\text{pop}() \cdot \alpha_3:\text{pop}() \cdot j$$

When verifying a stack implementation, we can determine if the correct linearization is $t_{1,2}$ or $t_{2,1}$ by inspecting the underlay state. If we focus on INVOKE and RETURN events, we can reconstruct the concrete trace t' , and if we focus on CCOMMIT and RCOMMIT, we can rebuild the linearized trace $t_{i,j}$.

Single possibilities are not sufficient in general when constructing proofs of future-dependent linearizations inductively. We must maintain multiple possibilities simultaneously to contend with current possibilities being invalidated in the future. To accomplish this, we lift the possibility axioms from a single possibility ρ to a non-empty *set* of possibilities $\bar{\rho}$, and define the lifted transition $\rightarrow$ as $\rightarrow$.³ In other words, a set of possibilities $\bar{\rho}$ can be updated into $\bar{\sigma}$ as long as every possibility in $\bar{\sigma}$ is reachable after zero or finitely many valid updates from some possibility in $\bar{\rho}$.

³ Formally, we have described how to construct the equivalent of forward simulations. Using a power-set construction, we obtain the equivalent of a forward-backward simulation, which is complete for trace equivalence [24].

In the stack example above, if the stack's linearizability proof were future-dependent (e.g., the timestamped stack [8]), we would maintain a set $\bar{\rho} = \{\rho_{1,2}, \rho_{2,1}\}$ for the trace t' . The possibility $\rho_{1,2}$ corresponds to the case $i = 1$ and $j = 2$, and $\rho_{2,1}$ the other case. We would then update $\bar{\rho}$ according to the updates in the example above for the corresponding assignments of i and j . When the `pop()` operation reaches its linearization point, we would find out which of the two pushes was deemed to have occurred first, and remove whichever of $\rho_{1,2}$ or $\rho_{2,1}$ has been invalidated.

5 Program Logic

Now we are ready to present our program logic (§ 5.1–§ 5.6) and its soundness and completeness (§ 5.7). Our program logic, building on that of [27], is a Rely-Guarantee program logic. The main idea of rely-guarantee reasoning [18] is that each thread over-approximates its local steps in terms of a *guarantee* relation $\mathcal{G}$, and over-approximates its environment steps (steps by other threads) in terms of a *rely* relation $\mathcal{R}$. One then shows that the thread guarantees its steps follow $\mathcal{G}$, while relying on its environment following $\mathcal{R}$. The main aspect of rely-guarantee reasoning is the notion of *stability*, which is what enables parallel composition to become a congruence.

Stability is the notion that when a thread shows that steps satisfy a predicate P , such as pre-conditions, invariants or post-conditions, these predicates must be shown to be stable with respect to $\mathcal{R}$, in the sense that:

$$\mathcal{R}; P \implies P \qquad P; \mathcal{R} \implies P$$

that is, predicates are invariant upon the environment taking steps *before* or *after* the predicates (where $-; -$ stands for relational composition). When a predicate satisfies this, we say it is *stable* w.r.t. $\mathcal{R}$, written $\text{Stable}(\mathcal{R}, P)$. Stability ensures that any predicates we verify a thread's local steps to satisfy will not be invalidated by having the scheduler introduce steps by other threads between local steps, and it is fundamental to ensuring threads may be appropriately parallel composed later.

Then, if a thread α is verified against $\mathcal{R}[\alpha]$ and $\mathcal{G}[\alpha]$, and β is verified against $\mathcal{R}[\beta]$ and $\mathcal{G}[\beta]$, we are allowed to parallel compose them when

$$\mathcal{G}[\alpha] \implies \mathcal{R}[\beta] \qquad \text{and} \qquad \mathcal{G}[\beta] \implies \mathcal{R}[\alpha]$$

which ensures that α and β follow each other's assumptions about their environments.

5.1 Proof State and Assertions

In LHL, a proof configuration is a tuple $(s, \bar{\rho})$ of a concrete LTS state $s \in \text{InterState}$ and a set of possibilities $\bar{\rho} \in \mathcal{P}(\text{Poss})$. The component s represents the current state of the program, which follows the actual semantics of the program, while $\bar{\rho}$ represents linearized overlay LTS states, which are managed by the prover following the possibility update rules in § 4.2.

By using a set of possibilities instead of a single possibility, users of LHL can track different linearizations corresponding to the current state. They can remove extra possibilities later when future events invalidate these linearizations. This allows us to support objects (such as the Herlihy-Wing queue [16]) with future-dependent linearizations [32], which is essential for the completeness proof of LHL.

We define preconditions P as predicates over the proof state, while rely/guarantee conditions $\mathcal{R}, \mathcal{G}$ and postconditions Q are defined as relations over the state. We use binary postconditions relating pre-/post-states to make certain proofs easier.

5.2 The Linearizability Hoare Logic

LHL consists of three judgments: (1) the module judgment $V_E, V_F, \mathcal{R}, \mathcal{G}, P, Q \models M$ groups all side conditions and implies our linearizability condition $(V_E \triangleright M_F) \rightsquigarrow V_F$, (2) the program judgment $\mathcal{R}, \mathcal{G} \vdash \{P\} e \{Q\}$ verifies the correctness of the method body, and (3) the commit judgment $\mathcal{G} \vdash_\alpha \{P\} m \{Q\}$ handles individual underlay events and applies possibility updates.

5.3 Module Judgment

To establish module correctness, the prover needs to find pre-/post-conditions for each $f \in F$ and rely/guarantee relations, and show the following:

$$\frac{\begin{array}{c} \forall \alpha, s, \bar{\rho}. (s, \bar{\rho}) \mathcal{R}[\alpha] (s, \bar{\rho}) \\ \forall \alpha. \mathcal{R}[\alpha]; \mathcal{R}[\alpha] \implies \mathcal{R}[\alpha] \quad \forall \alpha, \beta. \alpha \neq \beta \wedge \mathcal{G}[\alpha] \cup \text{Inv}[\alpha](-) \cup \text{Ret}[\alpha](-) \implies \mathcal{R}[\beta] \\ \forall \alpha, f. \text{Stable}(\mathcal{R}[\alpha], P[\alpha]_f) \quad \forall \alpha, f. P[\alpha]_f(\perp, \langle \text{init}(V_F), \perp, \perp \rangle) \\ \forall \alpha, f. \mathcal{R}[\alpha], \mathcal{G}[\alpha] \vdash \{P[\alpha]_f; \text{Inv}[\alpha](f)\} M[\alpha]^f \{Q[\alpha]_f\} \quad \forall \alpha, f, g. P[\alpha]_f; Q[\alpha]_f \implies P[\alpha]_g \end{array}}{V_E, V_F, \mathcal{R}, \mathcal{G}, P, Q \models M}$$

where $\text{Inv}(-) := \bigcup_{f \in F} \text{Inv}(f)$, $\text{Ret}(-) := \bigcup_{f \in F} \text{Ret}(f)$. The first few conditions are standard rely-guarantee conditions, respectively: that $\mathcal{R}[\alpha]$ is reflexive and self-stable, that threads respect each other's relies, and that method pre-conditions are stable. Other than the standard rely-guarantee side conditions, we require for all methods that (1) the initial proof state satisfies its pre-condition, (2) the required Hoare triples can be established for it, and (3) its post-condition implies the pre-condition of any method, so the system can still run safely afterwards.

► **Example 7** (Verifying the Exchanger Object). For the exchanger, we only need to find one pre-/post-condition for the **exch** method, which we made into an invariant I that holds at any step in all threads. It consists of three disjuncts that hold at different stages of the exchanger: when a thread makes an exchange offer, when the other thread accepts the offer, and when the offering thread clears the offer. They mainly relate the concrete state to the linearized state. For simplicity, we use a singleton possibility, as no future-dependent linearization is required here. For example, the **accepted** predicate requires the prover to linearize the response $\rho_R(\alpha)$ to the **exch** method made by the offering thread α when the CAS cell stores the **Acc** state.

$$\begin{aligned} I(s, \rho) &:= \exists \alpha, \beta, v, w. \text{offered}(\alpha, v)(s, \rho) \vee \text{accepted}(\alpha, \beta, v, w)(s, \rho) \vee \text{cleared}(s, \rho) \\ \text{offered}(\alpha, v)(s, \rho) &:= s_{\text{CAS}} = \text{Off}(\alpha, v) \wedge \rho_S = (\{\}, \{\}) \wedge \rho_C(\alpha) = \text{exch}(v) \wedge \rho_R(\alpha) = \perp \\ \text{accepted}(\alpha, \beta, v, w)(s, \rho) &:= s_{\text{CAS}} = \text{Acc}(\alpha, \beta, v, w) \wedge \rho_S = (\{\}, \{\}) \\ &\quad \wedge \rho_C(\alpha) = \text{exch}(v) \wedge \rho_R(\alpha) = \text{exch}(v).w \\ \text{cleared}(s, \rho) &:= s_{\text{CAS}} = \perp \wedge \rho_S = (\{\}, \{\}) \end{aligned}$$

The invariant obviously satisfies all conditions except the rely-guarantee-related ones and the Hoare triple, which we supply and prove later in this section.

5.4 Program Judgment

The program judgment consists of the following Hoare-style rules:

$$\begin{array}{c}
 \frac{P \implies Q \ v}{\mathcal{R}, \mathcal{G} \vdash \{P\} \text{Ret } v \{Q\}} \text{RET} \\
 \\
 \frac{\text{Stable}(\mathcal{R}, Q_S) \quad \mathcal{R}, \mathcal{G} \vdash \{P; Q_S\} p \{Q\} \quad P \xrightarrow{\alpha; \epsilon} \lambda s, t, \bar{\rho}. Q_S(s, \bar{\rho}, t, \bar{\rho}) \wedge \mathcal{G}(s, \bar{\rho}, t, \bar{\rho})}{\mathcal{R}, \mathcal{G} \vdash \{P\} \text{Tau } p \{Q\}} \text{SILENT} \\
 \\
 \frac{\mathcal{G} \vdash_{\alpha} \{P\} g(x) \{Q_C\} \quad \forall v. \mathcal{G} \vdash_{\alpha} \{P; Q_C\} g(x).v \{Q_R(v)\} \quad \text{Stable}(\mathcal{R}, Q_C) \quad \forall v. \text{Stable}(\mathcal{R}, Q_R(v)) \quad \forall v. \mathcal{R}, \mathcal{G} \vdash \{P; Q_C; Q_R(v)\} k(v) \{Q\}}{\mathcal{R}, \mathcal{G} \vdash \{P\} \text{Vis } g(x) k \{Q\}} \text{UNDERLAYSTEP}
 \end{array}$$

$$\text{where } P \xrightarrow{e} Q \iff \forall s, t, \bar{\rho}. P(s, \bar{\rho}) \wedge s \xrightarrow{e} t \Rightarrow Q(s, t, \bar{\rho})$$

These core rules are applicable to all programs. In practice, though, the user will define control structures along with their corresponding *derived rules*, and the core rules will not be used directly. For instance, `bind` and `call` may be defined, and their derived rules are as follows.

$$\begin{array}{c}
 \frac{\mathcal{R}, \mathcal{G} \vdash \{P\} e \{Q\} \quad \forall v. \mathcal{R}, \mathcal{G} \vdash \{Q(v)\} k(v) \{R\}}{\mathcal{R}, \mathcal{G} \vdash \{P\} v \leftarrow e; k(v) \{R\}} \\
 \\
 \frac{\text{Stable}(\mathcal{R}, Q) \quad \text{Stable}(\mathcal{R}, R) \quad \mathcal{G} \vdash_{\alpha} \{P\} g(x) \{Q\} \quad \forall v. \mathcal{G} \vdash_{\alpha} \{P; Q\} g(x).v \{R(v)\}}{\mathcal{R}, \mathcal{G} \vdash \{P\} g(x) \{P; Q; R\}}
 \end{array}$$

5.5 Commit Judgment

The final component of the proof is the *commit judgment* $\mathcal{G} \vdash_i \{P\} m \{Q\}$ generalizing linearization points into possibility updates.

$$\frac{P \xrightarrow{\alpha; m} \lambda s, t, \bar{\rho}. \exists \bar{\sigma}. \bar{\rho} \twoheadrightarrow^* \bar{\sigma} \wedge Q(s, \bar{\rho}, t, \bar{\sigma}) \wedge \mathcal{G}(s, \bar{\rho}, t, \bar{\sigma})}{\mathcal{G} \vdash_{\alpha} \{P\} m \{Q\}} \text{COMMIT}$$

The commit judgment asks the following: for any state $(s, \bar{\rho})$ satisfying the pre-condition P and any t reachable from s through the operational semantics, the user needs to produce new linearizations $\bar{\sigma}$ reachable after finitely many possibility updates ($\bar{\rho} \twoheadrightarrow^* \bar{\sigma}$, defined in § 4.4) and ensure the postcondition and guarantee relation are satisfied. In LHL, these updates are restricted to CCOMMIT or RCOMMIT; the necessary INVOKE and RETURN updates are applied automatically by the logic. The central proof obligation of LHL is to manage the set of possibilities $\bar{\rho}$ and maintain a sufficient set of correct linearizations. The user uses the commit judgment to update the linearization as they verify an operation, until they reach the postcondition Q_f . After this, the user will have to show that the series of commits they performed resulted in a linearization with the same return value as that produced by the operational semantics.

► **Example 8** (Verifying the Exchanger Object Cont.). We take line 5 in the exchanger code in Fig. 1 as an example:

$$\begin{array}{l}
 \{\text{offered}(\alpha, v)(s, \rho) \vee \exists \beta w. \text{accepted}(\alpha, \beta, v, w)(s, \rho)\} \\
 \text{if}(C.\text{cas}(\text{Off}(\alpha, v), \perp)) \{ \\
 \quad \left\{ \begin{array}{l} \text{cleared}(t, \sigma) \vee (\exists \beta w. \text{offered}(\beta, w)(t, \sigma) \\ \vee \exists \gamma u. \text{accepted}(\beta, \gamma, w, u)(t, \sigma)) \end{array} \right\} \wedge \rho_R(\alpha) = \text{exch}(v). \perp \} \\
 \}
 \end{array}$$

Here, we only consider the success branch of the CAS operation, and for simplicity, we only use a singleton possibility. Per the commit judgment, we need to prove:

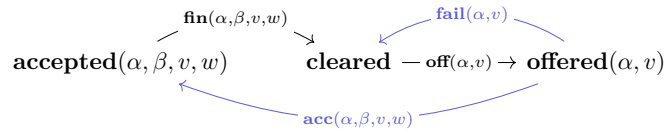
$$\begin{aligned} & \forall s, \rho, t. \left(\begin{array}{c} (\text{offered}(\alpha, v)(s, \rho) \vee \exists \beta w. \text{accepted}(\alpha, \beta, v, w)(s, \rho)) \\ \wedge s \xrightarrow{\alpha:\text{cas}(\text{Off}(\alpha, v)).\text{true}, \perp} t \end{array} \right) \\ & \Rightarrow \exists \sigma. \left(\begin{array}{c} \rho \xrightarrow{\alpha:\text{exch}(v), \alpha:\text{exch}(v).\perp} \sigma \wedge \mathcal{G}(s, \rho, t, \sigma) \wedge \\ (\text{cleared}(t, \sigma) \vee (\exists \beta w. \text{offered}(\beta, w)(t, \sigma) \\ \vee \exists \gamma u. \text{accepted}(\beta, \gamma, w, u)(t, \sigma))) \end{array} \right) \end{aligned}$$

We observe that the transition from s to t makes the **accepted** case of the precondition impossible, so we are left with **offered**(α, v) as the precondition. We choose to use $(\{\}, \{\})$ as σ (the same as ρ) and perform the possibility update ($\rho \dashrightarrow \sigma$) by making two transitions with event labels $\alpha:\text{exch}(v)$ and $\alpha:\text{exch}(v).\perp$. The return value of this linearization is $\perp$, and the actual return value of this branch is also $\perp$, so this linearization is valid. This possibility update is also safe because no concurrent exchange operation will successfully exchange its value for v and the new proof state satisfies the **cleared** disjunct of the post-condition.

Formally, the following rely-guarantee definition justifies this update, which matches the $\mathcal{G}_{\text{fail}}$ relation.

$$\begin{aligned} \mathcal{G}_{\text{fail}}(\alpha, v) (s, \rho, t, \sigma) &:= \left(\begin{array}{c} s \xrightarrow{\alpha:\text{cas}(a, \perp).\text{true}} t \\ \wedge \rho \xrightarrow{\alpha:\text{exch}(v), \alpha:\text{exch}(v).\perp} \sigma \end{array} \right) \\ \mathcal{G}_{\text{acc}}(\alpha, \beta, v, w) (s, \rho, t, \sigma) &:= \left(\begin{array}{c} s \xrightarrow{\beta:\text{cas}(\text{Off}(\alpha, v), \text{Acc}(\alpha, \beta, v, w)).\text{true}} t \\ \wedge \rho \xrightarrow{\alpha:\text{exch}(v), \beta:\text{exch}(w), \alpha:\text{exch}(v).w, \beta:\text{exch}(w).v} \sigma \end{array} \right) \\ \mathcal{G}[\alpha] &\triangleq \mathcal{G}_{\text{fail}}(\alpha, -) \cup \mathcal{G}_{\text{acc}}(\alpha, -, -, -) \cup \mathcal{G}_{\text{off}}(\alpha, -) \cup \mathcal{G}_{\text{fin}}(\alpha) \\ \mathcal{R}[\alpha] &\triangleq \cup_{\alpha' \neq \alpha} \mathcal{G}[\alpha'] \end{aligned}$$

Fig. 6 shows how these relations move the proof state among the three disjuncts of the invariant. After the current thread makes the offer, other threads can only perform the



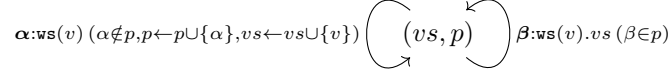
■ **Figure 6** Logical transitions used in the proof: each node represents the assertion satisfied by the proof state, and each edge is labelled with the subscript of the guarantee condition.

accept transition $\mathcal{G}_{\text{acc}}$, so the precondition in the Hoare triple is stable. The post-condition is also stable because no guarantee condition can step outside the invariant.

Other commands are verified similarly, and together, they establish the entire Hoare triple for the **exch** method:

$$\mathcal{R}[\alpha], \mathcal{G}[\alpha] \vdash \{I; \text{Inv}[\alpha]()\} M[\alpha]^{\text{exch}} \{I\}$$

which discharges the remaining obligation in the module judgment of the exchanger. We refer readers to the extended version and the Rocq mechanization for more details.



■ **Figure 10** Interval-sequential spec. of the write-snapshot.

5.6.2 Write-Snapshot

Snapshot objects are common in distributed systems – for instance, in implementing a consensus protocol. The particular write-snapshot implementation we verify [5] is our example of an interval linearizable object, which no existing system can verify. In addition, we are able to later use such objects, integrating them as components into other systems using vertical and horizontal composition. The specification of the write-snapshot is presented in Fig. 10.

In our system, interval-sequential specs have no restriction on the ordering of events. For instance, the spec in Fig. 10 validates the trace:

$$\alpha:\text{ws}(3) \cdot \beta:\text{ws}(5) \cdot \alpha:\text{ws}(3).\{3, 5\} \cdot \gamma:\text{ws}(2) \cdot \beta:\text{ws}(5).\{3, 5, 2\}$$

5.7 Soundness and Completeness

We prove our program logic to be both sound (Theorem 10) and complete (Theorem 11) with regard to compositional linearizability. We outline the proofs here.

► **Theorem 10 (Soundness).** *If $V_E, V_F, \mathcal{R}, \mathcal{G}, P, Q \models M$, then $V_E \triangleright M \rightsquigarrow V_F$.*

Proof. Recall that $V_E \triangleright M$ is a transition system which runs the code of M on top of V_E , i.e., there are four kinds of steps that can be made – an overlay call, an overlay return, an underlay call, or an underlay return. To show $V_E \triangleright M \rightsquigarrow V_F$, we can show that for every transition in $V_E \triangleright M$, there are zero or more corresponding transitions in V_F such that the two executions have the same overlay events. The module judgment and commit judgments directly provide the mappings from each individual step of $V_E \triangleright M$ to V_F , so for the soundness proof the task is to assemble those individual steps into a full execution.

We maintain that at all steps, for each thread α , the following invariants hold for the proof state $(s, \bar{p})$:

1. If the thread state of α is **Idle**, then for all $f \in F$, $P[\alpha]_f(s, \bar{p})$ holds.
2. If the thread state of α is **Cont** $f p$ for some $f \in F$, then there exists some I such that the following holds.

$$\begin{aligned} & \mathcal{R}[\alpha], \mathcal{G}[\alpha] \vdash \{P[\alpha]_f(s, \bar{p}); \text{Inv}(f); I\} p \{Q[\alpha]_f\} \\ & \wedge (P[\alpha]_f(s, \bar{p}); \text{Inv}(f); I)(s, \bar{p}) \wedge \text{Stable}(\mathcal{R}[\alpha]_f, I) \end{aligned}$$

3. If the thread state of α is **Ucall** $f g k$ for some $f \in F$ and $g \in E$, then there should exist some I and S such that the following holds.

$$\begin{aligned} & \mathcal{R}[\alpha], \mathcal{G}[\alpha] \vdash \{\mathcal{G}[\alpha]_f \vdash_\alpha \{P[\alpha]_f(s, \bar{p}); \text{Inv}(f); I\} g \{S\} \\ & \wedge P[\alpha]_f(s, \bar{p}); \text{Inv}(f); I; S\} p \{Q[\alpha]_f\} \wedge \text{Stable}(\mathcal{R}[\alpha]_f, I) \\ & \wedge \text{Stable}(\mathcal{R}[\alpha]_f, S) \wedge (P[\alpha]_f(s, \bar{p}); \text{Inv}(f); I)(s, \bar{p}) \end{aligned}$$

The proof then steps through the execution of $V_E \triangleright M$, cycling through the invariants in order for each thread.

For the overlay steps, we appeal to the precondition and postcondition of the program judgment. The module judgment requires the user to verify the operations with precondition $P[\alpha]_f$; $\text{Inv}(f)$ and postcondition $Q[\alpha]_f$. $\text{Inv}(f)$ is defined to make a simultaneous step in the operational semantics and the linearization, which takes care of that proof obligation. Finally, the module judgment has a side condition which asks the user to prove that a simultaneous overlay return and linearization return may be made if $Q[\alpha]_f$ holds of the state. This side-condition is present in the mechanization but omitted here, as it is almost always trivial but complicates the presentation.

For the underlay steps, recall that the commit judgment requires the user to directly provide linearization steps for the underlay event. Thus for this case we update I or supply S , and then simply add the provided steps to the linearization.

Once we assemble linearizations for the execution of each operation, we can chain those linearizations together to form a full linearization for the operational semantics of the entire concurrent object, proving $V_E \triangleright M \rightsquigarrow V_F$. ◀

► **Theorem 11 (Completeness).** *If $V_E \triangleright M \rightsquigarrow V_F$, then*

$$\exists \mathcal{R} \mathcal{G} P Q. V_E, V_F, \mathcal{R}, \mathcal{G}, P, Q \models M$$

Proof. For completeness, we note that the possibility update rules in Fig. 4 may be extended to be labelled by the same events as in Fig. 3. This extension is completely determined by the particular instantiation of the update rules. For example, if in **INVOKE**, we have that ρ_C goes to $\rho_C[\alpha \mapsto m]$, then the corresponding event is the overlay invocation $\alpha:m$. As a result, given a sequence of possibility updates q , we may take its projection $q \downarrow_{-,F}$ to the overlay events implied by the rule applications (**INVOKE** and **RETURN** rules), or its projection $q \downarrow_{E,-}$ to only underlay events (**CCOMMIT** or **RCommit**).

Completeness requires constructing $\mathcal{R}$, $\mathcal{G}$, P , and Q such that the program logic judgment holds. To do this, we define an invariant $I(s, \bar{\rho})$, which maintains that there exists a trace p (of the operational semantics in Fig. 3) such that the following three propositions hold.

1. The current state s is reachable by transitioning from the initial state to s along p following the operational semantics.
2. For every possibility ρ in $\bar{\rho}$, ρ is reachable through possibility updates q with the same overlay events as p (i.e., $q \downarrow_{-,F} = p \downarrow_{-,F}$).
3. For every possibility σ reachable by transitioning along the linearized specification on a trace q with the same overlay events as p (i.e., $q \downarrow_{-,F} = p \downarrow_{-,F}$), there is a prefix l of q such that a possibility ρ has transitioned along l , and there is a trace r such that ρ has transitioned along r into σ .

Informally, I maintains that the set of possibilities at any point contains exactly all possibilities that are still valid w.r.t. the current concrete trace, and that are reachable in the linearized specification. The rely-guarantee conditions are then defined using I as follows:

$$\begin{aligned} P(s, \bar{\rho}) &\iff I(s, \bar{\rho}) & Q(_, _, t, \bar{\sigma}) &\iff I(t, \bar{\sigma}) & \mathcal{R}(s, \bar{\rho}, t, \bar{\sigma}) &\iff I(s, \bar{\rho}) \Rightarrow I(t, \bar{\sigma}) \\ \mathcal{G}(_, _, t, \bar{\sigma}) &\iff I(t, \bar{\sigma}) \end{aligned}$$

From here we must prove the actual program logic judgments. The side-conditions on I , such as stability, are straightforward due to how I is defined. The key obligation of the program judgment is then to show that steps maintain the invariant. To ensure this, upon each commit, we advance the possibility set as far as possible, populating it with every possible linearization of the current execution that satisfies the invariant. Once we reach the end of the function, we perform the return step by again advancing the possibilities in

the set as far as possible, and then pruning all the possibilities which are not ready to make the overlay return. The assumption that the program is already linearizable is key here, as it allows us to prove that after the pruning is complete, the possibility set will still be non-empty, allowing us to complete the overlay return step. ◀

6 Related Work

6.1 Elimination-Backoff Stack and Write-Snapshot

Ours is the first paper with the technology to mechanize the proof of linearizability of the one-shot write-snapshot [5] object against its interval-sequential specification, and we believe we are also the first to give a truly modular proof of the EBStack. We compare other works that have verified the EBStack at the end of this section.

6.2 Possibility-Based Reasoning

Our approach is mainly inspired by [16], [21], and [27]. [16] not only proposed the notion of linearizability but also introduced possibility reasoning as an axiomatic method for proving atomic linearizability, and showed it is complete and sound for proofs of linearizability for individual traces. Their original technique has been recently mechanized by [17]. Their technique focuses only on atomicity, however, and only on verifying a single component, and does not provide the ability to verify large object systems either modularly or compositionally. Our approach is strictly more general in that it handles non-atomic objects and is compatible with theirs, so we could have verified any example they can verify.

[21] packaged possibility reasoning in a program logic for the first time, and used interval partial orders to represent a set of possibilities of an execution. [26, 27] show that [21]’s program logic is not complete for atomically linearizable objects, and extend it further to support compositional linearizability. They conjectured that their program logic was complete due to its relationship with possibility reasoning. Neither work has been mechanized. Our program logic mechanization builds upon their theory, and extends the possibilities from [27] with state, making possibility reasoning simpler and more usable than using either partial orders or traces, and embeds this method into a mechanized program logic. We provide a further extensive comparison with the closely related work of [27] in the extended version of the paper [14].

6.3 Other Works

It is worth mentioning that other than Herlihy-Wing possibilities, [28] have provided a technique based on forward and backward simulation that is sound and complete for Herlihy-Wing linearizability. Backward simulations are often deemed hard to reason about, however, because they go against the natural forward flow of executions, introduce backward non-determinism, require knowledge of the future, and often require introducing history variables to provide information that would be available under a forward technique. A similar approach is that of [6], who prove a direct simulation between the EB stack and its abstract model. However, this simulation is based on the execution of the entire thread pool. When showing the simulation, the prover needs to discuss the execution of all threads. In contrast, our approach and other program logics focus on enabling thread-local reasoning, which is more accessible to programmers and makes proofs more intuitive.

Regarding the mechanization, the closest work to ours is [22]. It focuses on using linearizability to verify serializability. Our encoding of modules as interaction trees is the same as theirs, and some aspects of their treatment of atomic linearizability are compatible

with ours. Despite that, their framework is restricted to atomicity. This means that their LTSs pair call and return events and cannot handle non-atomic objects. As a result, they can't compute overlay objects ($V \triangleright M$), as these can be non-atomic even if V is atomic, so they define an object as a pair of an underlay (atomic) LTS and a module instead. Meanwhile, an object for us is just an LTS. While they prove several compositional properties of the modules, which boil down to properties of interaction trees, we prove several more properties related to computing overlay objects, monotonicity w.r.t. trace refinement, etc. These enable the compositional linearizability framework and make up a significant portion of our mechanization. Due to the compatibility of their approach with ours, we believe that their results around encoding serializability as atomic linearizability could be carried over.

Although previous linearizability-based works cannot support non-atomic objects like the exchanger, it is still possible to use Hoare triples to give them a tight specification. [29] specifies the functional correctness of the exchange operation through Hoare triples and develops a Hoare logic verifying the exchanger in Rocq. However, due to the limitation of functional correctness, their Hoare triple specification can only be used when the execution of all interleaving threads is known, i.e., composed through the parallel composition rule. In contrast, our work focuses on object-level verification, where the client would access the object in an arbitrary concurrent environment. Moreover, [29] only verified a simple example as the client of the exchanger, instead of a complex object such as the elimination-backoff stack. It is unclear whether their logic can appropriately handle complex examples such as the helping mechanism in the elimination-backoff stack or future-dependent linearizations.

6.4 Logical Atomicity

Logical atomicity [7, 19] uses logically atomic Hoare triples $\langle P \rangle \text{op}() \langle Q \rangle$ to specify the ability of the given operation to perform the update abstracted as a transition from assertion P to assertion Q at a single physically atomic statement. Since its focus is on atomicity, it cannot generally specify set and interval linearizable objects satisfactorily. It is compositional as one may directly use the logically atomic Hoare triple in proofs of other functions that use this operation. It is known that a logically atomic triple implies Herlihy-Wing linearizability [2], but there is no evidence that logical atomicity can be used to specify all Herlihy-Wing linearizable objects, so its completeness is unknown. Moreover, logically atomic triples rely on a concrete assertion language, a concrete programming language, and their underlying semantic models. This limits the compositionality to within its own program logic framework, while our approach can specify any object using an abstract LTS, without assuming the actual semantic model or the program logic verifying the object against the specification.

6.5 Contextual Refinement

Contextual refinement provides an alternative to linearizability. It uses an abstract piece of code A_E to specify the concrete implementation M_E by requiring that the behavior of M_E running in any context C is bounded by the behavior of A_E running in the same context, usually denoted as $M_E \sqsubseteq A_E$. Contextual refinement is equivalent [9] to Herlihy-Wing linearizability and supports vertical composition in a sequential setting [30]. However, contextual refinement relies on a concrete programming language for the abstract code, and an embedding of this language into the original language for the concrete code is necessary for composing a concrete context with an “abstract” specification. In contrast, our framework only uses the signature of an object and applies to any programming language whose behaviors are characterized by call and return events, without any modification to the language. This

is a major benefit of verifying objects directly against linearizability specifications. Moreover, there has been no work to support modular verification of non-atomic objects like the exchanger and the interval-linearizable write-snapshot within the contextual refinement framework.

6.6 Relational Hoare logic

Relational Hoare logic [1] is widely used to show contextual refinement. [23] used local rely-guarantee reasoning to implement a relational logic that verifies the contextual refinement of concurrent objects w.r.t. their atomic abstract code. [35] later applied this technique to verify an operating system using their relational program logic. They achieved this using logically atomic triples [7] to establish a contextual refinement between the two. Their work allows modular verification, but with limited encapsulation: an object's abstract code A_F must also include all its sub-components' abstract code A_E . They can verify sub-components against corresponding abstract code A_E and use their Hoare triples to execute A_E embedded in A_F . Consequently, the top-level specification consists of very complicated abstract code that does not intuitively specify its behaviors. In contrast, our LTS specification encapsulates all details of the underlay object and provides an implementation-agnostic interface for users to verify overlay objects.

[31] encodes the contextual refinement obligation into a ghost state to verify it with CSL. Iris [19] is a higher-order logic for verifying fine-grained concurrent programs and also features logical atomicity. It leads to the ReLoC framework [10, 11, 33], which shows contextual refinement of concurrent programs using a program logic, with prophecy variables [20] enabling future-dependent linearization. ReLoC achieves compositionality by first encapsulating reusable code into smaller functions, then verifying these functions against logically atomic Hoare triples encoded as a logical relation, and lastly verifying the refinement of the top-level object using these intermediate proofs. However, it is unknown how to reuse the top-level refinement specification when verifying objects that use this top-level object. As a result, every time they build new objects using the one verified against the refinement specification, they need to prove a different functional correctness specification before they can reuse its proof. In contrast, our framework imposes LTS specifications uniformly at any level of abstraction.

As shown in Table 1, both lines of work have verified the elimination-backoff stack but failed to achieve the same level of compositionality as we do. [23] directly inlined all underlay code in the `EBStack` code. ReLoC [10] broke down the exchanger into three atomic operations: `make_offer`, `accept_offer`, and `revoke_offer`, exposing implementation details, and implemented the `EBStack` using these operations. They built and verified the `EBStack` using these three methods and their specifications. This effectively inlines a specific protocol for implementing the exchanger object within the stack implementation. As a result, to verify other objects using the exchanger, nothing except the specifications for these low-level atomic operations can be salvaged and reused from their proof. They still need to inline all the implementation code of the exchanger in the overlay object. In contrast, our approach allows direct composition above the ready-to-use exchanger implementation and specification.

7 Evaluation and Conclusion

Table 1 shows the evaluation of our program logic. There are mainly two other lines of work that mechanized the verification of the elimination-backoff stack. Neither can handle non-atomic objects, such as the exchanger and the elimination array, compositionally, and must embed them in their stack code, while LHL allows compositional verification of all three objects.

■ **Table 1** Evaluation of the elimination-backoff stack proof in LHL and other frameworks and lines of Rocq code of LHL. We use ○ for not verified, ● for verified but with sub-modules inlined, and ● for compositionally verified. The Elimination Array requires no lines of Spec because the Spec is the same as for the Exchanger. (*) Both R-LRG and ReLoC can verify the try stack as a standalone object but they chose not to in their formalization. (†) The elimination array uses the same specification as the exchanger.

		EBStack	Exchanger	Elimination Array	TryStack*
R-LRG[23]		●	○	○	○
ReLoC[11]		●	○	○	○
LHL		●	●	●	●
LOC of LHL in Rocq	Spec	50	38	0 [†]	38
	Code	26	28	4	21
	Proof	4056	4244	6230	3046
	Sound	Complete	Definitions & Lemmas		
		1077	412	812	

Although we have simplified possibility reasoning using LTS instead of partial orders and traces, the Rocq proof burden still turns out to be significant. As Table 1 indicates, the proofs are admittedly lengthy, reflecting the inherent complexity of reasoning about concurrent behavior. This is a consequence of a few factors. First, we did not design the program logic for ease of use, instead attempting to ensure that the proofs of soundness and completeness are easier to write. Nonetheless, a more usable program logic could be implemented and validated by translating its proofs into proofs in our more fine-grained logic. A second issue is that we have not implemented many tactics and auxiliary automation. We view proof automation as orthogonal to our contribution; our logic provides the semantic foundations upon which future automation can be built. The major difficulty we encounter in the soundness and completeness proofs is reasoning about execution traces. Even though we can hide traces from logic users using LTS, the internal proof still requires direct management of traces. We believe we can solve these issues with more automation, building on top of the foundation of LHL laid by this paper.

References

- 1 Nick Benton. Simple relational correctness proofs for static analyses and program transformations. *ACM SIGPLAN Notices*, 39(1):14–25, 2004. doi:10.1145/964001.964003.
- 2 Lars Birkedal, Thomas Dinsdale-Young, Armaël Guéneau, Guilhem Jaber, Kasper Svendsen, and Nikos Tzevelekos. Theorems for free from separation logic specifications. *Proceedings of the ACM on Programming Languages*, 5(ICFP):1–29, 2021. doi:10.1145/3473586.
- 3 Elizabeth Borowsky and Eli Gafni. Immediate atomic snapshots and fast renaming (extended abstract). In Jim Anderson and Sam Toueg, editors, *Proceedings of the Twelfth Annual ACM Symposium on Principles of Distributed Computing, Ithaca, New York, USA, August 15-18, 1993*, pages 41–51. ACM, 1993. doi:10.1145/164051.164056.
- 4 Jeremy S Bradbury, James R Cordy, and Juergen Dingel. Mutation operators for concurrent java (j2se 5.0). In *Second Workshop on Mutation Analysis (Mutation 2006-ISSRE Workshops 2006)*, pages 11–11. IEEE, 2006.
- 5 Armando Castañeda, Sergio Rajsbaum, and Michel Raynal. Specifying concurrent problems: Beyond linearizability and up to tasks. In *Proceedings of the 29th International Symposium on Distributed Computing - Volume 9363*, DISC 2015, pages 420–435, Berlin, Heidelberg, 2015. Springer-Verlag. doi:10.1007/978-3-662-48653-5_28.

- 6 Robert Colvin and Lindsay Groves. A scalable lock-free stack algorithm and its verification. In *Fifth IEEE International Conference on Software Engineering and Formal Methods (SEFM 2007)*, pages 339–348. IEEE, 2007. doi:10.1109/SEFM.2007.2.
- 7 Pedro da Rocha Pinto, Thomas Dinsdale-Young, and Philippa Gardner. Tada: A logic for time and data abstraction. In Richard E. Jones, editor, *ECOOP 2014 - Object-Oriented Programming - 28th European Conference, Uppsala, Sweden, July 28 - August 1, 2014. Proceedings*, volume 8586 of *Lecture Notes in Computer Science*, pages 207–231. Springer, Springer, 2014. doi:10.1007/978-3-662-44202-9_9.
- 8 Mike Dodds, Andreas Haas, and Christoph M. Kirsch. A scalable, correct time-stamped stack. In *Proceedings of the 42nd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, POPL '15, pages 233–246, New York, NY, USA, 2015. Association for Computing Machinery. doi:10.1145/2676726.2676963.
- 9 Ivana Filipovic, Peter W. O'Hearn, Noam Rinetzkzy, and Hongseok Yang. Abstraction for concurrent objects. In Giuseppe Castagna, editor, *Programming Languages and Systems, 18th European Symposium on Programming, ESOP 2009, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2009, York, UK, March 22-29, 2009. Proceedings*, volume 5502 of *Lecture Notes in Computer Science*, pages 252–266, Berlin, Heidelberg, 2009. Springer. doi:10.1007/978-3-642-00590-9_19.
- 10 Dan Frumin, Robbert Krebbers, and Lars Birkedal. Reloc: A mechanised relational logic for fine-grained concurrency. In Anuj Dawar and Erich Grädel, editors, *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2018, Oxford, UK, July 09-12, 2018*, pages 442–451. ACM, 2018. doi:10.1145/3209108.3209174.
- 11 Dan Frumin, Robbert Krebbers, and Lars Birkedal. Reloc reloaded: A mechanized relational logic for fine-grained concurrency and logical atomicity. *Log. Methods Comput. Sci.*, 17(3), 2021. doi:10.46298/lmcs-17(3:9)2021.
- 12 Éric Goubault, Jérémy Ledent, and Samuel Mimram. Concurrent specifications beyond linearizability. In Jiannong Cao, Faith Ellen, Luís Rodrigues, and Bernardo Ferreira, editors, *22nd International Conference on Principles of Distributed Systems, OPODIS 2018, Hong Kong, December 17-19, 2018*, volume 125 of *LIPIcs*, pages 28:1–28:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.OPODIS.2018.28.
- 13 Eashan Hatti, Arthur Oliveira Vale, Zhongye Wang, Yueyang Feng, and Zhong Shao. Linearizability Hoare Logic. Software, NSF 2019285, NSF 2313433, NSF 2442888, DARPA Agreement No. HR00112590130, swHId: swH:1:dir:ee421dd7fa7166ad053af08aa59a2cf82c194b33 (visited on 2026-06-10). URL: <https://github.com/ehatti/LHL/tree/ecoop-camera-ready>, doi:10.4230/artifacts.26388.
- 14 Eashan Hatti, Arthur Oliveira Vale, Zhongye Wang, Yueyang Feng, and Zhong Shao. A complete program logic for compositional linearizability. Technical Report YALEU/DCS/TR-1577, Yale University, 2026. URL: <https://flint.cs.yale.edu/publications/lhl.html>.
- 15 Danny Hendler, Nir Shavit, and Lena Yerushalmi. A scalable lock-free stack algorithm. In *Proceedings of the sixteenth annual ACM symposium on Parallelism in algorithms and architectures*, pages 206–215, 2004. doi:10.1145/1007912.1007944.
- 16 Maurice Herlihy and Jeannette M. Wing. Linearizability: A correctness condition for concurrent objects. *ACM Trans. Program. Lang. Syst.*, 12(3):463–492, July 1990. doi:10.1145/78969.78972.
- 17 Prasad Jayanti, Siddhartha Jayanti, Ugur Yavuz, and Lizzie Hernandez. A universal, sound, and complete forward reasoning technique for machine-verified proofs of linearizability. *Proc. ACM Program. Lang.*, 8(POPL), January 2024. doi:10.1145/3632924.
- 18 Cliff B. Jones. Tentative steps toward a development method for interfering programs. *ACM Trans. Program. Lang. Syst.*, 5(4):596–619, October 1983. doi:10.1145/69575.69577.

- 19 Ralf Jung, Robbert Krebbers, Jacques-Henri Jourdan, Aleš Bizjak, Lars Birkedal, and Derek Dreyer. Iris from the ground up: A modular foundation for higher-order concurrent separation logic. *Journal of Functional Programming*, 28:e20, 2018. doi:10.1017/S0956796818000151.
- 20 Ralf Jung, Rodolphe Lepigre, Gaurav Parthasarathy, Marianna Rapoport, Amin Timany, Derek Dreyer, and Bart Jacobs. The future is ours: prophecy variables in separation logic. *Proceedings of the ACM on Programming Languages*, 4(POPL):1–32, 2019. doi:10.1145/3371113.
- 21 Artem Khyzha, Mike Dodds, Alexey Gotsman, and Matthew Parkinson. Proving linearizability using partial orders. In *Programming Languages and Systems: 26th European Symposium on Programming, ESOP 2017, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2017, Uppsala, Sweden, April 22–29, 2017, Proceedings 26*, pages 639–667. Springer, 2017. doi:10.1007/978-3-662-54434-1_24.
- 22 Mohsen Lesani, Li-yao Xia, Anders Kaseorg, Christian J. Bell, Adam Chlipala, Benjamin C. Pierce, and Steve Zdancewic. C4: verified transactional objects. *Proc. ACM Program. Lang.*, 6(OOPSLA1), April 2022. doi:10.1145/3527324.
- 23 Hongjin Liang and Xinyu Feng. Modular verification of linearizability with non-fixed linearization points. In *Proceedings of the 34th ACM SIGPLAN Conference on Programming Language Design and Implementation*, pages 459–470, 2013. doi:10.1145/2491956.2462189.
- 24 Nancy A. Lynch and Frits W. Vaandrager. Forward and backward simulations, II: timing-based systems. *Inf. Comput.*, 128(1):1–25, July 1996. doi:10.1006/inco.1996.0060.
- 25 Gil Neiger. Set-linearizability. In *Proceedings of the Thirteenth Annual ACM Symposium on Principles of Distributed Computing*, PODC ’94, page 396, New York, NY, USA, 1994. Association for Computing Machinery. doi:10.1145/197917.198176.
- 26 Arthur Oliveira Vale, Zhong Shao, and Yixuan Chen. A compositional theory of linearizability. *Proc. ACM Program. Lang.*, 7(POPL), January 2023. doi:10.1145/3571231.
- 27 Arthur Oliveira Vale, Zhong Shao, and Yixuan Chen. A compositional theory of linearizability. *J. ACM*, 71(2), April 2024. doi:10.1145/3643668.
- 28 Gerhard Schellhorn, John Derrick, and Heike Wehrheim. A sound and complete proof technique for linearizability of concurrent data structures. *ACM Trans. Comput. Logic*, 15(4), September 2014. doi:10.1145/2629496.
- 29 Ilya Sergey, Aleksandar Nanevski, Anindya Banerjee, and Germán Andrés Delbianco. Hoare-style specifications as correctness conditions for non-linearizable concurrent objects. *ACM SIGPLAN Notices*, 51(10):92–110, 2016. doi:10.1145/2983990.2983999.
- 30 Youngju Song, Minki Cho, Dongjae Lee, Chung-Kil Hur, Michael Sammler, and Derek Dreyer. Conditional contextual refinement. *Proceedings of the ACM on Programming Languages*, 7(POPL):1121–1151, 2023. doi:10.1145/3571232.
- 31 Aaron Turon, Derek Dreyer, and Lars Birkedal. Unifying refinement and hoare-style reasoning in a logic for higher-order concurrency. In *Proceedings of the 18th ACM SIGPLAN international conference on Functional programming*, pages 377–390, 2013. doi:10.1145/2500365.2500600.
- 32 Viktor Vafeiadis. *Modular fine-grained concurrency verification*. PhD thesis, University of Cambridge, UK, July 2008. doi:10.48456/tr-726.
- 33 Simon Friis Vindum, Dan Frumin, and Lars Birkedal. Mechanized verification of a fine-grained concurrent queue from meta’s folly library. In *Proceedings of the 11th ACM SIGPLAN International Conference on Certified Programs and Proofs*, pages 100–115, 2022. doi:10.1145/3497775.3503689.
- 34 Li-yao Xia, Yannick Zakowski, Paul He, Chung-Kil Hur, Gregory Malecha, Benjamin C. Pierce, and Steve Zdancewic. Interaction trees: representing recursive and impure programs in Coq. *Proc. ACM Program. Lang.*, 4(POPL), December 2019. doi:10.1145/3371119.
- 35 Fengwei Xu, Ming Fu, Xinyu Feng, Xiaoran Zhang, Hui Zhang, and Zhaohui Li. A practical verification framework for preemptive os kernels. In *International Conference on Computer Aided Verification*, pages 59–79. Springer, 2016. doi:10.1007/978-3-319-41540-6_4.