Reasoning about Reconfigurations of Distributed Systems

Emma Ahrens*  Marius Bozga†  Radu Iosif‡

Joost-Pieter Katoen§

March 17, 2022

Abstract

This paper presents a Hoare-style calculus for formal reasoning about reconfiguration programs of distributed systems. Such programs create and delete components and/or interactions (connectors) while the system components change state according to their internal behaviour. Our proof calculus uses a resource logic, in the spirit of Separation Logic [61], to give local specifications of reconfiguration actions. Moreover, distributed systems with an unbounded number of components are described using inductively defined predicates. The correctness of reconfiguration programs relies on havoc invariants, that are assertions about the ongoing interactions in a part of the system that is not affected by the structural change caused by the reconfiguration. We present a proof system for such invariants in an assume/rely-guarantee style. We illustrate the feasibility of our approach by proving the correctness of real-life distributed systems with reconfigurable (self-adjustable) tree architectures.

1 Introduction

The relevance of dynamic reconfiguration. Dynamic reconfigurable distributed systems are used increasingly as critical parts of the infrastructure of our digital society, e.g. datacenters, e-banking and social networking. In order to address maintenance (e.g., replacement of faulty and obsolete network nodes by new ones) and data traffic issues (e.g., managing the traffic inside a datacenter [54]), the distributed systems community has recently put massive effort in designing algorithms for reconfigurable systems, whose network topologies change at runtime [36]. This development provides new impulses to distributed algorithm design [53] and has given rise to self-adjustable

*Emma.Ahrens@rwth-aachen.de, RWTH Aachen University, D-52056, Germany
†Marius.Bozga@univ-grenoble-alpes.fr, Univ. Grenoble Alpes, CNRS, Grenoble INP, VERIMAG, 38000, France
‡Radu.Iosif@univ-grenoble-alpes.fr, Univ. Grenoble Alpes, CNRS, Grenoble INP, VERIMAG, 38000, France
§katoen@cs.rwth-aachen.de, RWTH Aachen University, D-52056, Germany
network architectures whose topology reconfigurations are akin to amendments of dynamic data structures such as splay trees \[59\]. However, reconfiguration is an important source of bugs, that may result in denial of services or even data corruption\(^1\). This paper introduces a logical framework for reasoning about the safety properties of such systems, in order to prove e.g., absence of deadlocks or data races.

**Modeling distributed systems.** In this paper we model distributed systems at the level of abstraction commonly used in component-based design of large heterogenous systems \[51\]. We rely on a clean separation of (a finite-state abstraction of) the behaviour from the coordination of behaviors, described by complex graphs of components (nodes) and interactions (edges). Mastering the complexity of a distributed system requires a deep understanding of the coordination mechanisms. We distinguish between *endogenous* coordination, that explicitly uses synchronization primitives in the code describing the behavior of the components (e.g. semaphores, monitors, compare-and-swap, etc.) and *exogenous* coordination, that defines global rules describing how the components interact. These two orthogonal paradigms play different roles in the design of a system: exogenous coordination is used during high-level model building, whereas endogenous coordination is considered at a later stage of development, to implement the model using low-level synchronization primitives.

Here we focus on *exogenous coordination* of distributed systems, consisting of an unbounded number of interconnected components, with a flexible topology, i.e. not fixed \(\text{à priori}\). We abstract from low-level coordination mechanisms between processes such as semaphores, compare-and-swap operations and the like. Components behave according to a small set of finite-state abstractions of sequential programs, whose transitions are labeled with events. They communicate via interactions (handshaking) modeled as sets of events that occur simultaneously in multiple components. Despite their apparent simplicity, these models capture key aspects of distributed computing, such as message delays and transient faults due to packet loss. Moreover, the explicit graph representation of the network is essential for the modeling of dynamic reconfiguration actions.

**Programming reconfiguration** The study of dynamic reconfiguration has led to the development of a big variety of formalisms and approaches to specify the changes to the structure of a system using e.g., graph-based, logical or process-algebraic formalisms (see \[12\] and \[16\] for surveys). With respect to existing work, we consider a simple yet general imperative reconfiguration language, encompassing four primitive reconfiguration actions (creation and deletion of components and interactions) as well as non-deterministic reconfiguration triggers (constraints) evaluated on small parts of the structure and the state of the system. These features exist, in very similar forms, in the vast majority of existing graph-based reconfiguration formalisms e.g., using explicit reconfiguration scripts as in COMMUNITY \[70\], reconfiguration controllers expressed as production rules in graph-grammars \[50\], guarded reconfiguration actions in DR-BIP \[32\] and graph rewriting rules in REO \[45\], to cite only a few. In our model, the primitive reconfiguration actions are executed sequentially, but interleave with the firing of

\(^1\)E.g., Google reports a cloud failure caused by reconfiguration: https://status.cloud.google.com/incident/appengine/19007
interactions i.e., the normal execution of the system. Sequential reconfiguration is not a major restriction, as the majority of reconfiguration languages rely on a centralized management [12]. Nevertheless, for the sake of simplicity, most existing reconfiguration languages avoid the fine-grain interleaving of reconfiguration and execution steps i.e., they freeze the system’s execution during reconfiguration. Our choice of allowing this type of interleaving is more realistic and closer to real-life implementation. Finally, our language supports open reconfigurations, in which the number of possible configurations is unbounded [16], via non-deterministic choice and iteration.

An illustrative example. We illustrate the setting by a token ring example, consisting of a finite but unbounded number of components, indexed from 1 to \( n \), connected via an unidirectional ring (Fig. 1). A token may be passed from a component \( i \) in state \( T \) (it has a token) to its neighbour, with index \( (i \mod n) + 1 \), which must be in state \( H \) (it has a hole instead of a token). As result of this interaction, the \( i \)-th component moves to state \( H \) while the \( (i \mod n) + 1 \) component moves to state \( T \). Note that token passing interactions are possible as long as at least two components are in different states; if all the components are in the same state at the same time, the ring is in a deadlock configuration.

During operation, components can be added to, or removed from the ring. On removing the component with index \( i \), its incoming (from \( i - 1 \), if \( i > 1 \), or \( n \), if \( i = 1 \)) and outgoing (to \( (i \mod n) + 1 \)) connectors are deleted before the component is deleted, and its left and right neighbours are reconnected in order to re-establish the ring-shaped topology. Consider the program in Listing 1, where the variables \( x \), \( y \) and \( z \) are assigned indices \( i \), \( (i \mod n) + 1 \) and \( (i \mod n) + 2 \), respectively (assuming \( n > 2 \)). The program removes first the right connector between \( y \) and \( z \) (line 2), then removes the left connector between \( x \) and \( y \) (line 3), before removing the component indexed by \( y \) (line 4) and reconnecting the \( x \) and \( z \) components (line 5). Note that the order of the disconnect commands is crucial: assume that component \( x \) is the only one in state \( T \) in the entire system. Then the token may move from \( x \) to \( y \) and is deleted together with the component (line 4). In this case, the resulting ring has no token and the system is in a deadlock configuration. The reconfiguration program in Listing 2 is obtained by
swapping lines 2 and 3 from Listing 1. In this case, the deleted component is in state H before the reconfiguration and its left connector is removed before its right one, thus ensuring that the token does not move to the y component (deleted at line 4).

The framework developed in this paper allows to prove that e.g., when applied to a token ring with at least two components in state H and at least one component in state T, the program in Listing 2 yields a system with at least two components in different states, for any n > 2. Using, e.g. invariant synthesis methods similar to those described in [1, 22, 10, 11], an initially correct parametric systems can be automatically proved to be deadlock-free, after the application of a sequence of reconfiguration actions.

The contributions of this paper. Whereas various formalisms for modeling distributed systems support dynamic reconfiguration, the formal verification of system properties under reconfigurations has received scant attention. We provide a configuration logic that specifies the safe configurations of a distributed system. This logic is used to build Hoare-style proofs of correctness, by annotating reconfiguration programs (i.e. programs that delete and create interactions or components) with assertions that describe both the topology of the system (i.e. the components and connectors that form its coordinating architecture) and the local states of the components. The annotations of the reconfiguration program are proved to be valid under so-called havoc invariants, expressing global properties about the states of the components, that remain, moreover, unchanged under the ongoing interactions in the system. In order to prove these havoc invariants for networks of any size, we develop an induction-based proof system, that uses a parallel composition rule in the style of assume/rely-guarantee reasoning. In contrast with existing formal verification techniques, we do not consider the network topology to be fixed in advance, and allow it to change dynamically, as described by the reconfiguration program. This paper provides the details of our proof systems and the semantics of reconfiguration programs. We illustrate the usability of our approach by proving the correctness of self-adjustable tree architectures [62] and conclude with a list of technical problems relevant for the automation of our method.

Main challenges. Formal reasoning about reconfigurable distributed systems faces two technical challenges. The first issue is the huge complexity of nowadays distributed systems, that requires highly scalable proof techniques, which can only be achieved by local reasoning, a key ingredient of other successful proof techniques, based on Separation Logic [56]. To this end, atomic reconfiguration commands in our proof system are specified by axioms that only refer to the components directly involved in the action, while framing out the rest of the distributed system. This principle sounds appealing, but is technically challenging, as components from the local specification interfere with components from the frame\(^2\). To tackle this issue, we assume that frames are invariant under the exchange of messages between components (interactions) and discharge these invariance conditions using cyclic proofs. The inference rules used to write such proof rely on a compositional proof rule, in the spirit of rely/assume-guarantee reasoning [58, 42], whose assumptions about the environment behavior are automatically synthesized from the formulae describing the system and the environment.

\(^2\)Essentially the equivalent of the environment in a compositional proof system for parallel programs.
The second issue is dealing with the non-trivial interplay between reconfigurations and interactions. Reconfigurations change the system by adding/removing components/interactions while the system is running, i.e. while state changes occur within components by firing interactions. Although changes to the structure of the distributed system seem, at first sight, orthogonal to the state changes within components, the impact of a reconfiguration can be immense. For instance, deleting a component holding the token in a token-ring network yields a deadlocked system, while adding a component with a token could lead to a data race, in which two components access a shared resource simultaneously. Technically, this means that a frame rule cannot be directly applied to sequentially composed reconfigurations, as e.g. an arbitrary number of interactions may fire between two atomic reconfiguration actions. Instead, we must prove havoc invariance of the intermediate assertions in a sequential composition of reconfiguration actions. As an optimization of the proof technique, such costly checks do not have to be applied along sequential compositions of reconfiguration actions that only decrease/increase the size of the architecture; in such monotonic reconfiguration sequences, invariance of a set of configurations under interaction firing needs only to be checked in the beginning (for decreasing sequences) or in the end (for increasing sequences).

2 A Model of Distributed Systems

For a function \( f : A \rightarrow B \), we denote by \( \text{dom}(f) \) its domain and by \( f[a \leftarrow b] \) the function that maps \( a \) into \( b \) and behaves like \( f \) for all other elements from the domain of \( f \). By \( \text{pow}(A) \) we denote the powerset of a set \( A \). For a relation \( R \subseteq A \times A \), we denote by \( R^* \) its reflexive and transitive closure. Given sets \( A \) and \( B \), we write \( A \subseteq_{\text{fin}} B \) if \( A \) is a finite subset of \( B \) and define \( A \cup B \triangleq A \cup B \) if \( A \cap B = \emptyset \) and \( A \cup B \) is undefined, if \( A \cap B \neq \emptyset \).

We model a distributed system by a finite set \( C \subseteq_{\text{fin}} C \), where \( C \) is a countably infinite universe of components. The components in \( C \) are said to be present (in the system) and those from \( C \setminus C \) are absent (from the system).

The present components can be thought of as the nodes of a network, each executing a copy of the same program, called behavior in the following. The behavior is described by a finite-state machine \( B = (P,Q,\rightarrow) \), where \( P \) is a finite set of ports i.e., the event alphabet of the machine, \( Q \) is a finite set of states, and \( \rightarrow \subseteq Q \times P \times Q \) is a transition relation. We denote transitions as \( q \xrightarrow{p} q' \) instead of \( (q,p,q') \), the states \( q \) and \( q' \) being referred to as the pre- and post-state of the transition.

The network of the distributed system is described by a finite set \( I \subseteq_{\text{fin}} C \times P \times C \times P \) of interactions. Intuitively, an interaction \( (c_1,p_1,c_2,p_2) \) connects the port \( p_1 \) of component \( c_1 \) with the port \( p_2 \) of component \( c_2 \), provided that \( c_1 \) and \( c_2 \) are distinct components. Intuitively, an interaction \( (c_1,p_1,c_2,p_2) \) can be thought of as a joint execution of transitions labeled with the ports \( p_1 \) and \( p_2 \) from the components \( c_1 \) and \( c_2 \), respectively.

**Definition 1** A configuration is a quadruple \( \gamma = (C,I,\rho,\nu) \), where \( C \) and \( I \) describe the present components and the interactions of the system, \( \rho : C \rightarrow Q \) is a state map associating each present component a state of the common behavior \( B = (P,Q,\rightarrow) \).
and \( \nu : \mathcal{V} \rightarrow \mathbb{C} \) is a store that maps variables, taken from a countably infinite set \( \mathcal{V} \), to components (not necessarily present). We denote by \( \Gamma \) the set of configurations.

**Example 1** For instance, the configuration \((\mathcal{C}, \mathcal{I}, \rho, \nu)\) of the token ring system, depicted in Fig. 1 (left) has present components \( \mathcal{C} = \{c_1, \ldots, c_n\} \), interactions \( \mathcal{I} = \{(c_i, \text{out}, c_{(i + 1) \mod n}) \mid i \in [1, n]\} \) and state map given by \( \rho(c_1) = \top \) and \( \rho(c_i) = \bot \), for \( i \in [2, n] \). The store \( \nu \) is arbitrary.

Given a configuration \((\mathcal{C}, \mathcal{I}, \rho, \nu)\), an interaction \((c_1, p_1, c_2, p_2) \in \mathcal{I}\) is loose if and only if \( c_i \notin \mathcal{C} \), for some \( i = 1, 2 \). A configuration is loose if and only if it contains a loose interaction. Interactions (resp. configurations) that are not loose are said to be tight. In particular, loose configurations are useful for the definition of a composition operation, as the union of disjoint sets of components and interactions, respectively:

**Definition 2** The composition of two configurations \( \gamma_1 = (\mathcal{C}_1, \mathcal{I}_1, \rho_1, \nu) \), for \( i = 1, 2 \), is defined as \( \gamma_1 \circ \gamma_2 \overset{\text{def}}{=} (\mathcal{C}_1 \uplus \mathcal{C}_2, \mathcal{I}_1 \uplus \mathcal{I}_2, \rho_1 \uplus \rho_2, \nu) \). The composition \( \gamma_1 \circ \gamma_2 \) is undefined if either \( \mathcal{C}_1 \uplus \mathcal{C}_2 \) or \( \mathcal{I}_1 \uplus \mathcal{I}_2 \) is undefined. A composition \( \gamma_1 \circ \gamma_2 \) is trivial if \( \mathcal{C}_1 = \mathcal{I}_1 = \rho_1 = \emptyset \), for some \( i = 1, 2 \). A configuration \( \gamma_2 \) is a subconfiguration of \( \gamma_1 \), denoted \( \gamma_1 \subseteq \gamma_2 \), if and only if there exists a configuration \( \gamma_3 \in \Gamma \), such that \( \gamma_1 = \gamma_2 \circ \gamma_3 \).

Note that a tight configuration may be the result of composing two loose configurations, whereas the composition of tight configurations is always tight. The example below shows that, in most cases, a non-trivial decomposition of a tight configuration necessarily involves loose configurations.

**Example 2** Let \( \gamma = (\mathcal{C}_1, \mathcal{I}_1, \rho, \nu) \), where \( \mathcal{C}_1 = \{c_1\} \), \( \mathcal{I}_1 = \{(c_i, \text{out}, c_{(i + 3) \mod 3} + 1, \text{in})\} \), for all \( i \in [1, 3] \), \( \rho_1(c_1) = \rho_2(c_2) = \bot \) and \( \rho_3(c_3) = \top \). Then \( \gamma = \gamma_1 \circ \gamma_2 \circ \gamma_3 \) is the configuration from the top-left corner of Fig. 2, where the store \( \nu \) is arbitrary. Note that \( \gamma_1, \gamma_2, \) and \( \gamma_3 \) are loose, respectively, but \( \gamma \) is tight. Moreover, the only way of decomposing \( \gamma \) into two tight subconfigurations \( \gamma_1 ' \) and \( \gamma_2 ' \) is taking \( \gamma_1 ' = \gamma \) and \( \gamma_2 ' = (\emptyset, \emptyset, \emptyset, \nu) \), or viceversa.

A configuration is changed by two types of actions: (a) havoc actions change the local states of the components by executing interactions (that trigger simultaneous transitions in different components), without changing the structure or the store, and (b) reconfiguration actions that change the structure, store and possibly the state map of a configuration. We refer to Fig. 2 for a depiction of havoc and reconfiguration actions. Each havoc action is the result of executing a sequence of interactions (horizontally depicted using straight double arrows), whereas each reconfiguration action (vertically depicted using snake-shaped arrows) corresponds to a statement in a reconfiguration program. The two types of actions may interleave, yielding a transition graph with a finite but unbounded (parametric) or even infinite (obtained by iteratively adding new components) set of vertices (configurations).

---

\(^{3}\)Since \( \text{dom}(\rho_i) \subseteq \mathcal{C}_i \), for \( i = 1, 2 \) and \( \mathcal{C}_1 \cap \mathcal{C}_2 = \emptyset \), the disjointness condition is not necessary for state maps.
Figure 2: Havoc and Reconfigurations of a Token Ring

```
γ
```

that an action $\top$ the following rule:

The complete lattice $\pow(\Gamma) \triangleq \pow(\Gamma) \cup \{\top\}$. The complete lattice $(\pow(\Gamma), \subseteq, \cup, \cap)$ is extended with a greatest element $\top$, with the conventions $S \cup \top \triangleq \top$ and $S \cap \top \triangleq S$, for each $S \in \pow(\Gamma)$. We consider that an action $f$ is disabled in a configuration $\gamma$ if and only if $f(\gamma) = \emptyset$ and that it faults in $\gamma$ if and only if $f(\gamma) = \top$. Actions are naturally lifted to sets of configurations as $f(S) \triangleq \bigcup_{\gamma \in S} f(\gamma)$, for each $S \subseteq \Gamma$.

**Definition 3** The havoc action $h : \Gamma \to \pow(\Gamma)$ is defined as $h(\gamma) \triangleq \{\gamma' \mid \gamma \Rightarrow^* \gamma'\}$, where $\Rightarrow^*$ is the reflexive and transitive closure of the relation $\Rightarrow \subseteq \Gamma \times \Gamma$, defined by the following rule:

\[
\frac{(c_1, p_1, c_2, p_2) \in I \quad c_1, c_2 \in C \quad p(c_i) = q_i \quad q_i \xrightarrow{p_i} q_i'}{\text{(Havoc)}}
\]

Note that the havoc action is the result of executing any sequence of tight interactions, whereas loose interactions are simply ignored. The above definition can be generalized to multi-party interactions $(c_1, p_1, \ldots, c_n, p_n)$ with $n \geq 1$ pairwise distinct participant components $c_1, \ldots, c_n$, that fire simultaneously transitions of the behavior labeled with the ports $p_1, \ldots, p_n$, respectively. In particular, the interactions of arity $n = 1$ correspond to the local (silent) actions performed independently by a single component. To keep the presentation simple, we refrain from considering such generalizations, for the time being.

**Example 3** Let $\gamma_i = \{(c_1, c_2, c_3), \{c_i, out, c_i \text{ mod } 3 + 1, in\} \mid i \in [1, 3]\}, p_i, \nu)$, for $i \in [1, 3]$ be the top-most configurations from Fig. 2, where $p_1(c_1) = p_1(c_2) = H$, $p_1(c_3) = \top$, $p_2(c_1) = T$, $p_2(c_2) = p_2(c_3) = H$, $p_3(c_1) = p_3(c_3) = H$, $p_3(c_2) = \top$ and $\nu(x) = c_1$, $\nu(y) = c_2$, $\nu(z) = c_3$. Then $h(\gamma_i) = \{\gamma_1, \gamma_2, \gamma_3\}$, for all $i \in [1, 3]$. 

---

7
2.1 The Expressiveness of the Model

Before moving on with the definition of a logic describing sets of configurations (§3), a reconfiguration language and a proof system for reconfiguration programs (§4), we discuss the expressive power of the components-behavior-interactions model of distributed systems introduced so far, namely what kinds of distributed algorithms can be described in our model?

On one hand, this model can describe message-passing algorithms on networks with unrestricted topologies (pipelines, rings, stars, trees, grids, cliques, etc.), such as flooding/notification of a crowd, token-based mutual exclusion, deadlock problems (dining philosophers/cryptographers), etc. Moreover, the model captures asynchronous communication, via bounded message channels modeled using additional components\(^4\). Furthermore, transient faults (process delays, message losses, etc.) can be modeled as well, by nondeterministic transitions e.g., a channel component might chose to nondeterministically lose a message. In particular, having a single finite-state machine that describes the behavior of all components is not a limitation, because finitely many behaviors \(B_1, \ldots, B_m\) can be represented by state machines with disjoint transition graphs, the state map distinguishing between different behavior types – if \(\rho(c) = q\) and \(q\) is a state of \(B_i\), the value of \(\rho(c)\) can never change to a state of a different behavior \(B_j\), as the result of a havoc action.

On the other hand, the current model cannot describe complex distributed algorithms, such as leader election [21, 29], spanning tree [60, 46], topological linearization [39], Byzantine consensus [48] or Paxos parliament [47], due to the following limitations:

- Finite-state behavior is oblivious of the identity of the components (processes) in distributed systems of arbitrary sizes. For instance, there is no distributed algorithm over rings that can elect a leader under the assumption of anonymous processes [21, 29].

- Interactions between a bounded number of participants cannot describe broadcast between arbitrarily many components, as in most common consensus algorithms [48, 47].

We proceed in the rest of the paper under these simplifying assumptions (i.e., finite-state behavior and bounded-arity interactions), as our focus is modeling the reconfiguration aspect of a distributed system, and consider the following extensions for future work:

- **Identifiers in registers**: the behavior is described by a finite-state machine equipped with finitely many registers \(r\) holding component identifiers, that can be used to send \((p!r)\) and receive \((p?r)\) identifiers \((p\text{ stands for a port name})\), perform equality \((r = r')\) and strict inequality \((r < r')\) checks, with no other relation or function on the domain of identifiers. For instance, identifier-aware behaviors are considered in [2] in the context of bounded model checking i.e., verification \(^4\)The number of messages in transit depends on the number of states in the behavior; unbounded message queues would require an extension of the model to infinite-state behaviors.
of temporal properties (safety and liveness) under the assumption that the system has a ring topology and proceeds in a bounded number of rounds (a round is completed when every component has executed exactly one transition). Algorithms running on networks of arbitrary topologies (described by graphs) are modeled using distributed register automata [8], that offer a promising lead for verifying properties of distributed systems with mutable networks.

• Broadcast interactions: interactions involving an unbounded number of component-port pairs e.g., the $p_0$ ports of all components except for a bounded set $c_1, \ldots, c_k$, that interact with ports $p_1, \ldots, p_k$, for a given integer constant $k \geq 0$. Broadcast interactions are described using universal quantifiers in [10], where network topologies are specified using first-order logic. To accommodate broadcast interactions in our model, one has to redefine composition, by considering e.g., gluing of interactions, in addition to the disjoint union of configurations (Def. 2). Changing this definition would have a non-trivial impact on the configuration logic used to write assertions in reconfiguration proofs (§3).

Considering a richer model of behavior (e.g., register automata, timed automata, or even Markov decision processes) would impact mainly the part of the framework that deals with checking the properties (i.e., safety, liveness or havoc invariance) of a set even Markov decision processes) would impact mainly the part of the framework that deals with checking the properties (i.e., safety, liveness or havoc invariance) of a set of configurations described by a formula of the configuration logic (defined in §3) but should not, in principle, impact the configuration logic itself, the programming language or the proof system for reconfiguration programs (defined in §4). However, accommodating broadcast communication requires changes at the level of the logic and, consequently, the reconfiguration programs proof system.

3 A Logic of Configurations

We define a Configuration Logic (CL) that is, an assertion language describing sets of configurations. Let $\mathbb{A}$ be a countably infinite set of predicate symbols, where $\#(A) \geq 1$ denotes the arity of a predicate symbol $A \in \mathbb{A}$. The CL formulæ are inductively described by the following syntax:

$$\phi ::= \text{true} \mid \text{emp} \mid x = y \mid x \in q \mid \langle x_1, p_1, x_2, p_2 \rangle \mid A(x_1, \ldots, x_{\#(A)}) \mid \phi \land \phi \mid \phi \lor \phi \mid \neg \phi \mid \exists x. \phi$$

where $q \in Q$, $A \in \mathbb{A}$ are predicate symbols and $x, y, x_1, x_2, \ldots \in \mathbb{V}$ are variables. The atomic formulæ $x \in q$, $\langle x_1, p_1, x_2, p_2 \rangle$, and $A(x_1, \ldots, x_{\#(A)})$ are called component, interaction and predicate atoms, respectively. A formula is said to be quantifier-free if it has no occurrences of predicate atoms. By $\text{fv}(\phi)$ we denote the set of free variables in $\phi$, that do not occur within the scope of an existential quantifier. A formula is quantifier-free if it has no occurrence of existential quantifiers. A substitution is a partial mapping $\sigma : \mathbb{V} \rightarrow \mathbb{V}$ and the formula $\phi \sigma$ is the result of replacing each free variable $x \in \text{fv}(\phi) \cap \text{dom}(\sigma)$ by $\sigma(x)$ in $\phi$. We denote by $[x_1/y_1, \ldots, x_k/y_k]$ the substitution that replaces $x_i$ with $y_i$, for all $i \in [1, k]$. We use the shorthands $\text{false} \equiv \neg \text{true}$, $x \neq y \equiv \neg(x = y)$, $\phi_1 \lor \phi_2 \equiv \neg(\neg\phi_1 \land \neg\phi_2)$, $\forall x. \phi_1 \equiv \neg(\exists x. \neg\phi_1)$ and $x @\equiv \bigvee_{q \in Q} x @ q$. 


We distinguish the boolean (\(\land\)) from the separating (\(*\)) conjunction: \(\phi_1 \land \phi_2\) means that \(\phi_1\) and \(\phi_2\) hold for the same configuration, whereas \(\phi_1 \ast \phi_2\) means that \(\phi_1\) and \(\phi_2\) hold separately, on two disjoint parts of the same configuration. Intuitively, a formula \(\varphi\) describes empty configurations, with no components and interactions. \(\exists \{\}_q\) describes a configuration with a single component, given by the store value of \(x\), in state \(q\), and \(\{x_1,p_1,x_2,p_2\}\) describes a single interaction between ports \(p_1\) and \(p_2\) of the components given by the store values of \(x_1\) and \(x_2\), respectively. The formula \(x_1@q_1 \ast \ldots \ast x_n@q_n \ast \{x_1,p_1,x_2,p_2\} \ast \ldots \ast \{x_{n-1},p_{n-1},x_n,p_n\}\) describes a structure consisting of \(n\) pairwise distinct components, in states \(q_1,\ldots,q_n\), respectively, joined by interactions between ports \(p_i\) and \(p_{i+1}\), respectively, for all \(i \in [1,n-1]\).

The CL logic is used to describe configurations of distributed systems of unbounded size, by means of predicate symbols, defined inductively by a given set of rules. For reasons related to the existence of (least) fixed points, the definitions of predicates are symbolic configurations. The symbolic configurations are formulæ of the form \(\xi \land \pi\), where \(\xi\) and \(\pi\) are defined by the following syntax:

\[
\xi ::= \text{emp} \mid x@q \mid \{x_1,p_1,x_2,p_2\} \mid A(x_1,\ldots,x_{\#(A)}) \mid \xi \ast \xi \\
\pi ::= x = y \mid x \neq y \mid \pi \land \pi
\]

The interpretation of CL formulæ is given by a semantic relation \(\models_\Delta\), parameterized by a finite set of inductive definitions (SID) \(\Delta\), consisting of rules \(A(x_1,\ldots,x_{\#(A)}) \leftarrow \exists y_1 \ldots y_k . \phi\), where \(\phi\) is a symbolic configuration, such that \(\text{fv}(\phi) \subseteq \{x_1,\ldots,x_{\#(A)}\} \cup \{y_1,\ldots,y_k\}\). The relation \(\models_\Delta\) is defined inductively on the structure of formulæ, as follows:

\[
\begin{align*}
(C, I, \rho, v) \models_\Delta \text{true} & \iff C = \emptyset \text{ and } I = \emptyset \\
(C, I, \rho, v) \models_\Delta \text{emp} & \iff C = \emptyset \\
(C, I, \rho, v) \models_\Delta x = y & \iff v(x) = v(y) \\
(C, I, \rho, v) \models_\Delta x @ q & \iff C = \{v(x)\}, I = \emptyset \text{ and } \rho(v(x)) = q \\
(C, I, \rho, v) \models_\Delta \{x_1,p_1,x_2,p_2\} & \iff C = \emptyset, I = \{(v(x_1),p_1,v(x_2),p_2)\} \\
(C, I, \rho, v) \models_\Delta A(x_1,\ldots,x_{\#(A)}) & \iff (C, I, \rho, v) \models_\Delta \phi[x_1/y_1,\ldots,x_{\#(A)}/y_{\#(A)}], \text{ for some rule } A(x_1,\ldots,x_{\#(A)}) \models_\Delta \phi \text{ from } \Delta \\
(C, I, \rho, v) \models_\Delta \phi_1 \ast \phi_2 & \iff \text{there exist configurations } \gamma_1 \text{ and } \gamma_2, \text{ such that } (C, I, \rho, v) = \gamma_1 \ast \gamma_2 \text{ and } \gamma_i \models_\Delta \phi_i \text{, for both } i = 1,2 \\
(C, I, \rho, v) \models_\Delta \neg \phi_1 & \iff (C, I, \rho, v) \not\models_\Delta \phi_1 \\
(C, I, \rho, v) \models_\Delta \exists x . \phi_1 & \iff (C, I, \rho, v[x \leftarrow c],\rho) \models_\Delta \phi_1, \text{ for some } c \in C
\end{align*}
\]

From now on, we consider the SID \(\Delta\) to be clear from the context and write \(\gamma \models \phi\) instead of \(\gamma \models_\Delta \phi\). If \(\gamma \models \phi\), we say that \(\gamma\) is a model of \(\phi\) and define the set of models of \(\phi\) as \[[\phi]\] \(= \{\gamma \mid \gamma \models \phi\}\). A formula \(\phi\) is satisfiable if and only if \([\phi] \neq \emptyset\). Given formulæ \(\phi\) and \(\psi\), we say that \(\phi\) entails \(\psi\) if and only if \([\phi] \subseteq [\psi]\), written \(\phi \models \psi\).

**Example 4** The SID below defines chains of components and interactions, with at least \(h,t \in \mathbb{N}\) components in state \(H\) and \(T\), respectively:

\[
\begin{align*}
\text{chain}_{0,1}(x,x) & \leftarrow x@T \\
\text{chain}_{h,0}(x,y) & \leftarrow \exists \xi. x@T \ast (x.out, z.in) \ast \text{chain}_{h-1,t}(z,y) \\
\text{chain}_{1,0}(x,x) & \leftarrow x@H \\
\text{chain}_{h,t}(x,y) & \leftarrow \exists \xi. x@H \ast (x.out, z.in) \ast \text{chain}_{h-1,t}(z,y) \\
\text{chain}_{0,0}(x,x) & \leftarrow x@-
\end{align*}
\]
where \( k \geq 1 \) is defined as \( \max(k - 1, 0) \), for all \( k \in \mathbb{N} \). The configurations \( \{c_1, \ldots, c_n\} \) \( \{c_i, \text{out}, c_i \mod n \} \) for all \( i \in [1,n] \), \( \rho, v \) from Example 1 are models of the formula \( \exists x \exists y . \text{chain}_{0,0}(x,y) \land (y.\text{out},x.\text{in}) \), for all \( n \in \mathbb{N} \). This is because any such configuration can be decomposed into a model of \( (y.\text{out},x.\text{in}) \) and a model of \( \text{chain}_{0,0}(x,y) \). The latter is either a model of \( \text{out} \) matching the body of the rule \( \text{chain}_{0,0}(x,y) \leftarrow x@\_ \) if \( x = y \), or a model of \( \exists z . \text{out} \) * \( \langle x.\text{out},z.\text{in} \rangle \text{chain}_{0,0}(z,y) \), matching the body of the rule \( \text{chain}_{0,0}(z,y) \leftarrow \exists z . \text{out} \) * \( \langle x.\text{out},z.\text{in} \rangle \text{chain}_{0,0}(z,y) \).

### 3.1 The Expressiveness of CL

The CL logic is quite expressive, due to the interplay between first-order quantifiers and inductively defined predicates. For instance, the class of cliques, in which there is an interaction between the \textit{out} and \textit{in} ports of any two present components are defined as by the formula:

\[
\forall x \forall y . \text{out}(x,y) \rightarrow \text{true} \rightarrow (\langle x.\text{out},y.\text{in} \rangle) \text{true}
\]

Describing clique-structured networks is important for modeling consensus protocols, such as Byzantine [48] or Paxos [47].

\textit{Connected} networks, used in e.g., linearization algorithms [38], are such that there exists a path of interactions between each two present components in the system:

\[
\forall x \forall y . \text{out}(x,y) \rightarrow \text{true} \rightarrow \text{reach}(x,y)
\]

where the predicate \( \text{reach}(x,y) \) is defined by the following rules:

\[
\text{reach}(x,y) \leftarrow x = y, \quad \text{reach}(x,y) \leftarrow \exists z . \text{out}(x,z) \rightarrow \text{true}
\]

A grid is a connected network that, moreover, satisfies the following formula:

\[
\forall x \forall y \forall z . \langle x.\text{out},y.\text{in} \rangle \rightarrow \langle x.\text{out},z.\text{in} \rangle \rightarrow \text{true} \rightarrow \exists u . \langle y.\text{out},u.\text{in} \rangle \rightarrow \langle z.\text{out},u.\text{in} \rangle \rightarrow \text{present}(x,y,z,u)
\]

where \( \text{present}(x_1,\ldots,x_n) \overset{\text{def}}{=} \forall 1 \leq i \leq n . x_i@\_ \) states that the components \( x_1,\ldots,x_n \) are present and pairwise distinct. Grids are important in modeling distributed scientific computing [37].

As suggested by work on Separation Logic [26], the price to pay for this expressiveness is the inherent impossibility of having decision procedures for a fragment of CL, that combines first-order quantifiers with inductively defined predicates. A non-trivial fragment of CL that has decision procedures for satisfiability and entailment is the class of symbolic configurations [9]. However, we do not expect to describe systems with clique or grid network topologies using symbolic configurations. Furthermore, we conjecture that these classes are beyond the expressiveness of the symbolic configuration fragment. We discuss these issues in more detail in §7.

### 4 A Language for Programming Reconfigurations

This section defines \textit{reconfiguration} actions that change the structure of a configuration. We distinguish between reconfigurations and havoc actions (Def. 3), that change configurations in orthogonal ways (see Fig. 2 for an illustration of the interplay between
the two types of actions). The reconfiguration actions are the result of executing a
given reconfiguration program on the distributed system at hand. This section presents
the syntax and operational semantics of the reconfiguration language. Later on, we in-
troduce a Hoare-style proof system to reason about the correctness of reconfiguration
programs.

4.1 Syntax and Operational Semantics

Reconfiguration programs, ranged over by \( R \), are inductively defined by the following
syntax:

\[
R ::= \text{new}(q,x) \mid \text{delete}(x) \mid \text{connect}(x_1,p_1,x_2,p_2) \mid \text{disconnect}(x_1,p_1,x_2,p_2)
\mid \text{with } x_1,\ldots,x_k : \theta \text{ do } R_1 \od \mid R_1; R_2 \mid R_1 + R_2 \mid R_1^* 
\]

where \( q \in Q \) is a state, \( x,x_1,x_2,\ldots \in \mathbb{V} \) are program variables and \( \theta \) is a predicate-free
quantifier-free formula of the CL logic, called a trigger.

The primitive commands are \( \text{new}(q,x) \) and \( \text{delete}(x) \), that create and delete a com-
ponent (the newly created component is set to execute from state \( q \)) given by the store
value of \( x \), \( \text{connect}(x_1,p_1,x_2,p_2) \) and \( \text{disconnect}(x_1,p_1,x_2,p_2) \), that create and delete
an interaction, between the ports \( p_1 \) and \( p_2 \) of the components given by the store values
of \( x_1 \) and \( x_2 \), respectively. We denote by \( \mathcal{P} \) the set of primitive commands.

A conditional is a program of the form \( \text{with } x_1,\ldots,x_k : \theta \text{ do } R \od \) that performs
the following steps, \textit{with no havoc action (Def. 3) in between the first and second steps}
below:

1. maps the variables \( x_1,\ldots,x_k \) to some components \( c_1,\ldots,c_k \in C \) such that the
configuration after the assignment contains a model of the trigger \( \theta \); the condi-
tional is disabled if the current configuration is not a model of \( \exists x_1 \vdots \exists x_k . \theta \) *
true,
2. launches the first command of the program \( R \) on this configuration, and
3. continues with the remainder of \( R \), in interleaving with havoc actions;
4. upon completion of \( R \), the values of \( x_1,\ldots,x_k \) are forgotten.

To avoid technical complications, we assume that nested conditionals use pairwise dis-
joint tuples of variables; every program can be statically changed to meet this condi-
tion, by renaming variables. Note that the trigger \( \theta \) of a conditional \( \text{with } x_1,\ldots,x_k : \theta \text{ do } R \od \) has no quantifiers nor predicate atoms, which means that the overall number of
components and interactions in a model of \( \theta \) is polynomially bounded by the size
of \( \text{(number of symbols needed to represent)} \theta \). Intuitively, this means that the part of
the system (matched by \( \theta \)) to which the reconfiguration is applied is relatively small,
thus the procedure that evaluates the trigger can be easily implemented in a distributed
environment, as e.g., consensus between a small number of neighbouring components.

The sequential composition \( R_1; R_2 \) executes \( R_1 \) followed by \( R_2 \), with zero or more
interactions firing in between. This is because, even though being sequential, a recon-
figuration program runs in parallel with the state changes that occur as a result of firing
more times in sequence, nondeterministically.

As a matter of fact, the conditionals are the only constraint. This design choice sustains the view of a distributed system as a cloud of components and interactions in which reconfigurations can occur anywhere a local condition is met. In other words, we do not need variable assignments to traverse the architecture — the program works rather by identifying a part of the system that matches a small pattern, and applying the reconfiguration locally to that subsystem. For instance, a typical pattern for writing reconfiguration programs is (with \( x_1 : \theta_1 \) do \( R_1 \) od + ... + with \( x_k : \theta_k \) do \( R_k \) od \( )^* \), where \( R_1, \ldots, R_k \) are loop-free sequential compositions of primitive commands. This program continuously choses a reconfiguration sequence \( R \) nondeterministically and either applies it on a small part of the configuration that satisfies \( \theta_i \), or does nothing, if no such subconfiguration exists within the current configuration.

The operational semantics of reconfiguration programs is given by the structural rules in Fig. 3, that define the judgements \( R : \gamma \rightarrow \gamma' \) and \( \gamma \rightarrow \gamma' \), where \( \gamma \) and \( \gamma' \) are configurations and \( R \) is a program. Intuitively, \( R : \gamma \rightarrow \gamma' \) means that \( \gamma' \) is a successor of \( \gamma \) following the execution of \( R \) and \( \gamma \rightarrow \gamma' \) means that \( R \) faults in \( \gamma \). The semantics of

---

**Figure 3: Operational Semantics of the Reconfiguration Language**

| Rule | Definition |
|------|------------|
| new\((q,x)\) | \( (C, I, p, v) \rightarrow (C \cup \{c\}, I, p[e \leftarrow q], v[x \leftarrow c]) \) |
| delete\((x)\) | \( (C, I, p, v) \rightarrow (C \setminus \{v(x)\}, I, p, v) \) if \( v(x) \in C \) and \( v(x) \notin C \) if \( v(x) \notin C \) |
| fulfill \( (x, \gamma) \) | \( \varphi \rightarrow \top \) if \( \gamma \in \text{init} \) and \( \gamma \rightarrow \top \) if \( \varphi \rightarrow \top \)
| \( c \in C \setminus \mathcal{C} \) | \( \text{new}(q, x) : (C, I, p, v) \rightarrow (C \cup \{c\}, I, p[e \leftarrow q], v[x \leftarrow c]) \)
| delete\((x)\) | \( (C, I, p, v) \rightarrow (C \setminus \{v(x)\}, I, p, v) \)
| connect\((x_1, p_1, x_2, p_2)\) | \( (C, I, p, v) \rightarrow (C \cup \{(v(x_1), p_1, v(x_2), p_2)\}, p, v) \)
| disconnect\((x_1, p_1, x_2, p_2)\) | \( (C, I, p, v) \rightarrow (C \setminus \{(v(x_1), p_1, v(x_2), p_2)\}, p, v) \)

with \( x_1, \ldots, x_k : \varphi \) do \( \text{R od} : (C, I, p, v) \rightarrow (C', I', p', v' \{x_1 \leftarrow c_1', \ldots, x_k \leftarrow c_k'\}) \)

\[
\begin{align*}
R_1 : \gamma & \rightarrow \gamma' \quad R_2 : \gamma & \rightarrow \gamma' \\
R_1 ; R_2 : \gamma & \rightarrow \gamma' \\
R_1 + R_2 : \gamma & \rightarrow \gamma'
\end{align*}
\]

\[
\begin{align*}
R^n : \gamma & \rightarrow \gamma' \\
R^* & \rightarrow \gamma' \\
R^* & \rightarrow \gamma'
\end{align*}
\]

the interactions. Last, \( R_1 + R_2 \) executes either \( R_1 \) or \( R_2 \), and \( R^* \) executes \( R \) zero or more times in sequence, nondeterministically.

It is worth pointing out that the reconfiguration language does not have explicit assignments between variables. As a matter of fact, the conditionals are the only constructs that nondeterministically bind variables to indices that satisfy a given logical constraint. This design choice sustains the view of a distributed system as a cloud of components and interactions in which reconfigurations can occur anywhere a local condition is met. In other words, we do not need variable assignments to traverse the architecture — the program works rather by identifying a part of the system that matches a small pattern, and applying the reconfiguration locally to that subsystem. For instance, a typical pattern for writing reconfiguration programs is (with \( x_1 : \theta_1 \) do \( R_1 \) od + ... + with \( x_k : \theta_k \) do \( R_k \) od \( )^* \), where \( R_1, \ldots, R_k \) are loop-free sequential compositions of primitive commands. This program continuously choses a reconfiguration sequence \( R \) nondeterministically and either applies it on a small part of the configuration that satisfies \( \theta_i \), or does nothing, if no such subconfiguration exists within the current configuration.
a program $R$ is the action $\langle\langle R \rangle\rangle: \Gamma \rightarrow \text{pow}(\Gamma)^\top$, defined as:

$$
\langle\langle R \rangle\rangle(\gamma) \overset{\text{def}}{=} \begin{cases} 
\top & \text{if } R : \gamma \Rightarrow \gamma' \\
\{ & \gamma' \mid R : \gamma \Rightarrow \gamma' \} & \text{otherwise}
\end{cases}
$$

The only primitive commands that may fault are $\text{delete}(x)$ and $\text{disconnect}(x_1.p_1,x_2.p_2)$; for both, the premisses of the faulty rules are disjoint from the ones for normal termination, thus the action $\langle\langle R \rangle\rangle$ is properly defined for all programs $R$. Notice that the rule for sequential composition uses the havoc action $h$ in the premiss, thus capturing the interleaving of havoc state changes and reconfiguration actions.

### 4.2 Reconfiguration Proof System

To reason about the correctness properties of reconfiguration programs, we introduce a Hoare-style proof system consisting of a set of axioms that formalize the primitive commands (Fig. 4a), a set of inference rules for the composite programs (Fig. 4b) and a set of structural rules (Fig. 4c). The judgements are Hoare triples $\{ \phi \} R \{ \psi \}$, where $\phi$ and $\psi$ (called pre- and postcondition, respectively) are CL formulae. The triple $\{ \phi \} R \{ \psi \}$ is valid, written $\models \{ \phi \} R \{ \psi \}$, if and only if $\langle\langle R \rangle\rangle([\phi]) \subseteq [\psi]$. Note that a triple is valid only if the program does not fault on any model of the precondition. In other words, an invalid Hoare triple $\{ \phi \} R \{ \psi \}$ cannot distinguish between $\langle\langle R \rangle\rangle([\phi]) \not\subseteq [\psi]$ (non-faulting incorrectness) and $\langle\langle R \rangle\rangle([\psi]) = \top$ (faulting).

The axioms (Fig. 4a) give the local specifications of the primitive commands in the language by Hoare triples whose preconditions describe only those resources (components and interactions) necessary to avoid faulting. In particular, $\text{delete}(x)$ and $\text{disconnect}(x_1.p_1,x_2.p_2)$ require a single component $x@$ and an interaction $\langle x_1.p_1,x_2.p_2 \rangle$ to avoid faulting, respectively. The rules for sequential composition and iteration (Fig 4b) use the following semantic side condition, based on the havoc action (Def. 3):

**Definition 4** A formula $\phi$ is havoc invariant if and only if $h([\phi]) \subseteq [\phi]$.

Note that the dual inclusion $[\phi] \subseteq h([\phi])$ always holds, because $h$ is the reflexive and transitive closure of the $\Rightarrow$ relation (Def. 3). Since havoc invariance is required to prove the validity of Hoare triples involving sequential composition, it is important to have a way of checking havoc invariance. We describe a proof system for such havoc queries in §5. Moreover, the side condition of the consequence rule (Fig. 4c left) consists of two entailments, that are discharged by an external decision procedure (discussed in §7).

The frame rule (Fig. 4c bottom-right) allows to apply the specification of a local program, defined below, to a set of configurations that may contain more resources (components and interactions) than the ones asserted by the precondition. Intuitively, a local program requires a bounded amount of components and interactions to avoid faulting and, moreover, it only changes the configuration of the local subsystem, not affecting the entire system’s configuration. Formally, the set $\mathcal{L}$ of local programs is the least set that contains the primitive commands $\mathcal{P}$ and is closed under the application of the following rules:

$$
R \in \mathcal{L} \Rightarrow \text{with } x : \pi \text{ do } R \od \in \mathcal{L}, \text{ if } \pi \text{ is a conjunction of (dis-)equalities } \quad R_1,R_2 \in \mathcal{L} \Rightarrow R_1 + R_2 \in \mathcal{L}
$$
Lemma 1

For every program \( R \in \mathcal{L} \), the action \( \langle R \rangle \) is local for \( \text{modif}(R) \).
Moreover, $\mathcal{L}$ is precisely the set of programs with local semantics, as conditionals and sequential compositions (hence also iterations) are not local, in general:

**Example 5** To understand why $\mathcal{L}$ is precisely the set of local commands, consider the programs:

- (with $x : x@q$ do delete(x) od) is not local because, letting $\gamma_1$ be a configuration with zero components and $\gamma_2$ be a configuration with one component in state $q$, we have:
  \[
  \langle\langle \text{with } x : x@q \text{ do delete(x) od} \rangle \rangle (\gamma_1 \bullet \gamma_2) = \langle\langle \text{with } x : x@q \text{ do delete(x) od} \rangle \rangle (\gamma_1) \bullet \{\gamma_2\} = \emptyset \bullet \{\gamma_2\} = \emptyset.
  \]

- (skip; skip) is not local because, considering the system from Fig. 1, if we take $\gamma_1$ and $\gamma_2$, such that $\gamma_1 = x@T$ and $\gamma_2 = \langle\langle x.out, y.in \rangle\rangle * y@H$, we have:
  \[
  \langle\langle \text{skip; skip} \rangle \rangle (\gamma_1 \bullet \gamma_2) = [\langle\langle x@T * (x.out, y.in) * y@H \rangle\rangle] \cup [\langle\langle x@H * (x.out, y.in) * y@T \rangle\rangle]
  \]
  whereas $\langle\langle \text{skip; skip} \rangle \rangle (\gamma_1) \bullet \{\gamma_2\} = [\langle\langle x@T * (x.out, y.in) * y@H \rangle\rangle]$.

We write $\vdash \{\phi\} \triangleright \{\psi\}$ if and only if $\{\phi\} \triangleright \{\psi\}$ can be derived from the axioms using the inference rules from Fig. 4 and show the soundness of the proof system in the following. The next lemma gives sufficient conditions for the soundness of the axioms (Fig. 4a):

**Lemma 2** For each axiom $\{\phi\} \triangleright \{\psi\}$, where $R \in \mathcal{P}$ is primitive, we have $\langle\langle R \rangle\rangle(\langle\langle \phi \rangle\rangle) = \langle\langle \psi \rangle\rangle$.

The soundness of the proof system in Fig. 4 follows from the soundness of each inference rule:

**Theorem 1** For any Hoare triple $\{\phi\} \triangleright \{\psi\}$, if $\vdash \{\phi\} \triangleright \{\psi\}$ then $\models \{\phi\} \triangleright \{\psi\}$.

As an optimization, reconfiguration proofs can often be simplified, by safely skipping the check of one or more havoc invariant side conditions of sequential compositions, as explained below.

**Definition 6** A program of the form disconnect($x_1,p_1,x'_1,p'_1$); ... disconnect($x_k,p_k,x'_k,p'_k$); connect($x_{k+1},p_{k+1},x'_{k+1},p'_{k+1}$); ... connect($x_{\ell},p_{\ell},x'_{\ell},p'_{\ell}$) is said to be a single reversal program.

Single reversal programs first disconnect components and then reconnect them in a different way. For such programs, only the first and last application of the sequential composition rule require checking havoc invariance:

**Proposition 1** Let $R = \text{disconnect}(x_1,p_1,x'_1,p'_1); \ldots \text{disconnect}(x_k,p_k,x'_k,p'_k); \text{connect}(x_{k+1},p_{k+1},x'_{k+1},p'_{k+1}); \ldots \text{connect}(x_{\ell},p_{\ell},x'_{\ell},p'_{\ell})$ be a single reversal program. If $\phi_0, \ldots, \phi_\ell$ are CL formulae, such that:

- $\models \{\phi_{i-1}\} \triangle \{x_i,p_i,x'_i,p'_i\} \{\phi_i\}$, for all $i \in [1,k]$,
- $\models \{\phi_{j-1}\} \triangle \{x_j,p_j,x'_j,p'_j\} \{\phi_j\}$, for all $j \in [k+1,\ell]$, and
- $\phi_1$ and $\phi_{\ell-1}$ are havoc invariant,

then we have $\models \{\phi_0\} \triangleright \{\phi_\ell\}$.
4.3 Examples of Reconfiguration Proofs

We prove that the outcome of the reconfiguration program from Fig. 1 (Listing 2), started in a token ring configuration with at least two components in state H and at least one in state T, is a token ring with at least one component in each state. The pre- and postcondition are $\exists x \exists y . \text{chain}_{2,1}(x,y) \ast (\langle y . \text{out}, x . \text{in} \rangle)$ and $\exists x \exists y . \text{chain}_{1,1}(x,y) \ast (\langle y . \text{out}, x . \text{in} \rangle)$, respectively, with the definitions of $\text{chain}_{h,t}(x,y)$ given in Example 4, for all constants $h,t \in \mathbb{N}$.

\[
\{ \exists x \exists y . \text{chain}_{2,1}(x,y) \ast (\langle y . \text{out}, x . \text{in} \rangle) \}
\]

with $x,y,z : (\langle \text{out}, \text{in} \rangle) \ast y @ H \ast (\langle \text{out}, \text{in} \rangle)$ do

\[
\{ (\exists x \exists y . \text{chain}_{2,1}(x,y) \ast (\langle y . \text{out}, x . \text{in} \rangle)) \land (\langle \text{out}, \text{in} \rangle) \ast y @ H \ast (\langle \text{out}, \text{in} \rangle) \ast \text{true} \} \quad (*)
\]

\[
\{ (\langle \text{out}, \text{in} \rangle) \ast y @ H \ast (\langle \text{out}, \text{in} \rangle) \ast \text{true} \}
\]

\[
\{ \text{disconnect}(x, \text{out}, y . \text{in}) ; \text{disconnect}(y, \text{out}, z . \text{in}) ; \}
\]

\[
\{ \text{delete}(y) ; \}
\]

\[
\{ \text{connect}(x, \text{out}, z . \text{in}) \}
\]

\[
\{ \text{chain}_{1,1}(z,x) \}
\]

\[
\{ \forall \exists y . \text{chain}_{1,1}(x,y) \ast (\langle y . \text{out}, x . \text{in} \rangle) \}
\]

The inference rule for conditional programs sets up the precondition (*) for the body of the conditional. This formula is equivalent to $(\langle \text{out}, \text{in} \rangle) \ast y @ H \ast (\langle \text{out}, \text{in} \rangle) \ast \text{true}$.

We have considered the reconfiguration program from Fig. 1 (Listing 2) which deletes a component from a token ring. The dual operation is the addition of a new component. Here the precondition states that the system is a valid token ring, with at least one component in state H and at least another one in state T. We prove that the execution of the dual program yields a token ring with at least two components in state H, as the new component is added without a token.
5 The Havoc Proof System

This section describes a set of axioms and inference rules for proving the validity of havoc invariance queries of the form \( h(\llbracket \phi \rrbracket) \subseteq \llbracket \phi \rrbracket \), where \( \phi \) is a CL formula interpreted over a given SID and \( h \) is the havoc action (Def. 3). Such a query is valid if and only if the result of applying any sequence of interactions on a model of \( \phi \) is again a model of \( \phi \) (Def. 4). Havoc invariance queries occur as side conditions in the rules for sequential composition and iteration (Fig. 4b) of reconfiguration programs. Thus, having a proof system for havoc invariance is crucial for the applicability of the rules in Fig. 4 to obtain proofs of reconfiguration programs.

The havoc proof system uses a compositional rule, able to split a query of the form \( h(\llbracket \phi_1 \star \phi_2 \rrbracket) \subseteq \llbracket \psi_1 \star \psi_2 \rrbracket \) into two queries \( h(\llbracket \phi_i \star \psi_i \rrbracket) \subseteq \llbracket \psi_i \rrbracket \), where each frontier formula \( \psi_i \) defines a set of interactions that over-approximate the effect of executing the system described by \( \phi_i \) (resp. \( \psi_i \)) over the one described by \( \phi_i \) (resp. \( \psi_i \)), for \( i = 1, 2 \). In principle, the frontier formulae \( (\psi_1 \text{ and } \psi_2) \) can be understood as describing the interference between parallel actions in an assume/rely guarantee-style parallel composition rule [58, 42]. In particular, since the frontier formulae only describe interactions and carry no state information whatsoever, such assumptions about events triggered by the environment are reminiscent of compositional reasoning about input/output automata [23].

Compositional reasoning about havoc actions requires the following relaxation of the definition of havoc state changes (Def. 3), by allowing the firing of loose, in addition to tight interactions:

**Definition 7** The following rules define a relation \( (c_1,p_1,c_2,p_2) \subseteq \Gamma \times \Gamma \), parameterized by a given interaction \( (c_1,p_1,c_2,p_2) \):

\[
(\text{Loose})\quad (c_1,p_1,c_2,p_2) \in I \quad c_i \in C, c_{3-i} \not\in C \quad \rho(c_i) = q_i \quad q_i \xrightarrow{p_i} q_i' \quad i = 1, 2 \\
\quad (C, I, \rho, v) \xrightarrow{(c_1,p_1,c_2,p_2)} (C, I, \rho[c_i \leftarrow q_i'], v)
\]

\[
(\text{Tight})\quad (c_1,p_1,c_2,p_2) \in I \quad c_1 \neq c_2 \in C \quad \rho(c_i) = q_i \quad q_i \xrightarrow{p_i} q_i' \quad i = 1, 2 \\
\quad (C, I, \rho, v) \xrightarrow{(c_1,p_1,c_2,p_2)} (C, I, \rho[c_1 \leftarrow q_1'][c_2 \leftarrow q_2'], v)
\]

For a sequence \( w = i_1 \ldots i_n \) of interactions, we define \( \rightarrow^w \) to be the composition of \( \rightarrow, \ldots, \rightarrow \), assumed to be the identity relation, if \( w \) is empty.
The difference with Def. 3 is that only the states of the components from the configuration are changed according to the transitions in the behavior. This more relaxed definition matches the intuition of partial systems in which certain interactions may be controlled by an external environment; those interactions are conservatively assumed to fire anytime they are enabled by the components of the current structure, independently of the environment.

Example 6 (contd. from Example 3) Let \( \tilde{\gamma}_i = \{c_2, c_3\}, \{(c_1, \text{out}, c_1 \mod 3 + 1, \text{in}) \mid i \in [1, 3]\}, \rho, \nu \), for \( i \in [1, 3] \) be the top-most configurations from Fig. 2 without the \( c_1 \) component, where \( \rho_1(c_2) = H, \rho_1(c_3) = T, \rho_2(c_2) = \rho_2(c_3) = H, \rho_3(c_2) = T, \rho_3(c_3) = H \). Then, by executing the loose interactions \( (c_3, \text{out}, c_1, \text{in}) \) and \( (c_1, \text{out}, c_2, \text{in}) \) from \( \tilde{\gamma}_1 \), we obtain:

\[
\begin{align*}
\tilde{\gamma}_1 \xrightarrow{(c_3, \text{out}, c_1, \text{in})} \tilde{\gamma}_2 \xrightarrow{(c_1, \text{out}, c_2, \text{in})} \tilde{\gamma}_3
\end{align*}
\]

Executing the tight interaction \( (c_2, \text{out}, c_3, \text{in}) \) from \( \tilde{\gamma}_3 \) leads back to \( \tilde{\gamma}_1 \) i.e., \( \tilde{\gamma}_3 \xrightarrow{(c_2, \text{out}, c_3, \text{in})} \tilde{\gamma}_1. \)

### 5.1 Regular Expressions

Proving the validity of a havoc query \( \mathbf{h}(\emptyset) \subseteq \mathbf{\psi} \) involves reasoning about the sequences of interactions that define the outcome of the havoc action. We specify languages of such sequences using extended regular expressions, defined inductively by the following syntax:

\[
L ::= \varepsilon | \Sigma[\alpha] | L \cdot L | L \cup L | L^* \mid L \searrow \eta \mid L
\]

where \( \varepsilon \) denotes the empty string, \( \Sigma[\alpha] \) is an alphabet symbol associated with either an interaction atom or a predicate atom \( \alpha \) and \( \cdot, \cup \) and \( ^* \) are the usual concatenation, union and Kleene star. By \( L_1 \searrow_{\eta_1, \eta_2} L_2 \) we denote the interleaving (zip) product of the languages described by \( L_1 \) and \( L_2 \) with respect to the sets \( \eta_1 \) and \( \eta_2 \) of alphabet symbols of the form \( \Sigma[\alpha] \), respectively.

The language of a regular expression \( L \) in a configuration \( \eta = (\mathcal{C}, \iota, \rho, \nu) \) is defined below:

\[
\begin{align*}
&\langle \varepsilon \rangle(\eta) \overset{\text{def}}{=} \{ \varepsilon \} & &\langle \Sigma[\alpha] \rangle(\eta) \overset{\text{def}}{=} \bigcup \{ I \mid (\mathcal{C}, \iota, \rho, \nu) \subseteq \gamma, (\mathcal{C}, \iota, \rho, \nu) \models \alpha \} \\
&\langle L_1 \cdot L_2 \rangle(\eta) \overset{\text{def}}{=} \{ w_1w_2 \mid w_i \in \langle L_i \rangle(\eta), i = 1, 2 \} & &\langle L_1 \cup L_2 \rangle(\eta) \overset{\text{def}}{=} \langle L_1 \rangle(\eta) \cup \langle L_2 \rangle(\eta) \\
&\langle L^* \rangle(\eta) \overset{\text{def}}{=} \bigcup_{i \geq 0} \langle L^i \rangle(\eta) & &\langle L \searrow \eta \rangle(\eta) \overset{\text{def}}{=} \{ w \mid w \searrow_{\eta_1, \eta_2} \in \langle L_i \rangle(\eta), i = 1, 2 \}
\end{align*}
\]

where \( \langle \eta \rangle(\gamma) \overset{\text{def}}{=} \bigcup_{\Sigma[\alpha] \in \eta} \langle \Sigma[\alpha] \rangle(\gamma) \) and \( w \searrow_{\eta}(\gamma) \) is the word obtained from \( w \) by deleting each symbol not in \( \langle \eta \rangle(\gamma) \) from it. The \( i \)-th composition of \( L \) with itself is defined, as usual, by \( L^0 = \varepsilon \) and \( L^{i+1} = L^i \cdot L, \) for \( i \geq 0 \). We denote by \( \text{supp}(L) \) the support of \( L \) i.e., set of alphabet symbols \( \Sigma[\alpha] \) from the regular expression \( L \).

Example 7 Let \( \eta = (\{c_1, c_2, c_3, c_4\}, \{(c_1, \text{out}, c_2, \text{in}), (c_2, \text{out}, c_3, \text{in}), (c_3, \text{out}, c_4, \text{in})\}, \rho, \nu) \) be a configuration, such that \( \nu(x) = c_1 \), \( \nu(y) = c_2 \) and \( \nu(z) = c_3 \). Then, we have

\[
\begin{align*}
&\langle \Sigma[x, \text{out}, y, \text{in}] \rangle(\eta) = \{(c_1, \text{out}, c_2, \text{in})\}, \langle \Sigma[y, \text{out}, z, \text{in}] \rangle(\eta) = \{(c_2, \text{out}, c_3, \text{in})\} \quad \text{and} \\
&\langle \Sigma[\text{chain}_0, 0, (x, \text{z})] \rangle(\eta) = \{(c_1, \text{out}, c_2, \text{in}), (c_2, \text{out}, c_3, \text{in})\}.
\end{align*}
\]
Given a configuration $\gamma$ and a predicate atom $\alpha$, there can be, in principle, more than one subconfiguration $\gamma' \subseteq \gamma$ such that $\gamma' \models \alpha$. This is problematic, because then $\langle \Sigma[\alpha] \rangle(\gamma)$ may contain interactions from different subconfigurations of $\gamma$ that are models of $\alpha$, thus cluttering the definition of the language $\langle \Sigma[\alpha] \rangle(\gamma)$. We fix this issue by adapting the notion of precision, originally introduced for SL [18, 57], to our configuration logic:

**Definition 8 (Precision)** A formula $\phi$ is precise on a set $S$ of configurations if and only if, for every configuration $\gamma \in S$, there exists at most one configuration $\gamma'$, such that $\gamma' \subseteq \gamma$ and $\gamma' \models \phi$. A set of formulæ $\Phi$ is precisely closed if $\psi$ is precise on $\llbracket \psi \rrbracket$, for any two formulæ $\phi, \psi \in \Phi$.

Symbolic configurations using predicate atoms are not precise for $\Gamma$, in general\(^5\). To understand this point, consider a configuration consisting of two overlapping models of $\text{chain}_{h,t}(x,y)$, starting and ending in $x$ and $y$, respectively, with a component that branches on two interactions after $x$ and another component that joins the two branches before $y$. Then $\text{chain}_{h,t}(x,y)$ is not precise on such configurations (that are not models of $\text{chain}_{h,t}(x,y)$ whatsoever). On the positive side, we can state the following:

**Proposition 2** The set of symbolic configurations built using predicate atoms $\text{chain}_{h,t}(x,y)$, for $h, t \geq 0$ (Example 4) is precisely closed.

Two regular expressions are congruent if they denote the same language, whenever interpreted in the same configuration. Lifted to models of a symbolic configuration, we define:

**Definition 9** Given a symbolic configuration $\phi$, the regular expressions $L_1$ and $L_2$ are congruent for $\phi$, denoted $L_1 \cong_\phi L_2$, if and only if $\llbracket L_1 \rrbracket(\gamma) = \llbracket L_2 \rrbracket(\gamma)$, for all configurations $\gamma \in \llbracket \phi \rrbracket$.

Despite the universal condition that ranges over a possibly infinite set of configurations, congruence of regular expressions with alphabet symbols of the form $\Sigma[\alpha]$, where $\alpha$ is an interaction or a predicate atom, is effectively decidable by an argument similar to the one used to prove equivalence of symbolic automata [25].

### 5.2 Inference Rules for Havoc Triples

We use judgements of the form $\eta \triangleright \llbracket \phi \rrbracket L \llbracket \psi \rrbracket$, called havoc triples, where $\phi$ and $\psi$ are SL formulæ, $L$ is a regular expression, and $\eta$ is an environment (a set of alphabet symbols), whose role will be made clear below (Def. 14 and Lemma 3). A havoc triple states that each finite sequence of (possibly loose) interactions described by a word in $L$, when executed in a model of the precondition $\phi$, yields a model of the postcondition $\psi$.

**Definition 10** A havoc triple $\eta \triangleright \llbracket \phi \rrbracket L \llbracket \psi \rrbracket$ is valid, written $\models \eta \triangleright \llbracket \phi \rrbracket L \llbracket \psi \rrbracket$ if and only if, for each configuration $\gamma \in \llbracket \phi \rrbracket$, each sequence of interactions $w \in \llbracket L \rrbracket(\gamma)$ and each configuration $\gamma'$, such that $\gamma \xrightarrow{w} \gamma'$, we have $\gamma' \in \llbracket \psi \rrbracket$.

\(^5\)Unlike the predicates that define acyclic data structures (lists, trees) in SL, which are typically precise.
For a symbolic configuration \( \phi \), we denote by \( \text{inter}(\phi) \) and \( \text{preds}(\phi) \) the sets of interaction and predicate atoms from \( \phi \), respectively and define the set of atoms \( \text{atoms}(\phi) \equiv \text{inter}(\phi) \cup \text{preds}(\phi) \) and the regular expression \( \Sigma[\phi] \equiv \bigcup_{x \in \text{atoms}(\phi)} \Sigma[x] \). We show that the validity of a havoc triple is a sufficient argument for the validity of a havoc query; because havoc triples are evaluated via open state changes (Def. 10), the dual implication is not true, in general.

**Proposition 3** If \( \models \eta \triangleright \{[\phi]\} \Sigma[\phi]^* \{[\psi]\} \), then \( h([\phi]) \subseteq [\psi] \).

We describe next a set of axioms and inference rules used to prove the validity of havoc triples. For a symbolic configuration \( \phi \), we write \( x \simeq_\emptyset y \) (\( x \not\simeq_\emptyset y \)) if and only if the equality (disequality) between \( x \) and \( y \) is asserted by the symbolic configuration \( \phi \), e.g., \( x \simeq_{\text{emp}} x = z \vdash x = y \) and \( x \not\simeq_{\text{emp}} x \not\simeq_{\text{emp}} y \); note that \( x \not\simeq_\emptyset y \) is not necessarily the negation of \( x \simeq_\emptyset y \).

**Definition 11** For a symbolic configuration \( \phi \) and an interaction atom \( \langle x_1.p_1.x_2.p_2 \rangle \), we write:

- \( \phi \uparrow \langle x_1.p_1.x_2.p_2 \rangle \) if and only if \( \phi \) contains a subformula \( y@q \), such that \( y \simeq_\emptyset x_i \) and \( q \) is not the pre-state of some behavior transition with label \( p_i \), for some \( i = 1, 2 \); intuitively, any interaction defined by the formula \( \langle x_1.p_1.x_2.p_2 \rangle \) is disabled in any model of \( \phi \),

- \( \phi \downarrow \langle x_1.p_1.x_2.p_2 \rangle \) if and only if, for each interaction atom \( \langle y_1.p'_1.y_2.p'_2 \rangle \in \text{inter}(\phi) \), there exists \( i \in [1, 2] \), such that \( x_i \not\simeq_\emptyset y_i \); intuitively, the interaction defined by the formula \( \langle x_1.p_1.x_2.p_2 \rangle \) is not already present in a model of \( \phi \); i.e., \( \langle x_1.p_1.x_2.p_2 \rangle \not\in \phi \) is satisfiable.

The axioms (Fig. 5a) discharge valid havoc triples for the empty sequence (\( \emptyset \)), that changes nothing and the sequence consisting of a single interaction atom, that can be either disabled in every model (\( \uparrow \)), or enabled in some model (\( \Sigma \)) of the precondition, respectively; in particular, the (\( \Sigma \)) axiom describes the open state change produced by an interaction (Def. 7), firing on a (possibly empty) set of components, whose states match the pre-states of transitions for the associated behaviors. The (\( \perp \)) axiom discharges trivially valid triples with unsatisfiable (false) preconditions.

The redundancy rule (\( \leftarrow \)) in Fig. 5b removes an interaction atom from the precondition of a havoc triple, provided that the atom is never interpreted as an interaction from the language denoted by the regular expression from the triple. Conversely, the rule (\( \rightarrow \)) adds an interaction to the precondition, provided that the precondition (with that interaction atom) is consistent. Note that, without the \( \phi \uparrow \alpha \) side condition, we would obtain a trivial proof for any triple, by adding an interaction atom twice to the precondition, i.e. using the rule (\( \rightarrow \)), followed by (\( \perp \)).

The composition rule (\( \triangleright \)) splits a proof obligation into two simpler havoc triples (Fig. 5c). The pre- and postconditions of the premisses are subformule of the pre- and postcondition of the conclusion, joined by separating conjunction and extended by so-called frontier formulæ, describing those sets of interaction atoms that may cross the boundary between the two separated conjuncts. The frontier formulæ play the role of
Figure 5: Proof System for Havoc Triples

\[
\begin{align*}
(\varepsilon) & \quad \frac{\eta \vdash \{\emptyset\} \cap \{\emptyset\}}{\eta \vdash \{\emptyset\} \cap \{\emptyset\}} \\
(\top) & \quad \frac{\eta \vdash \{\emptyset\} \Sigma[\alpha]\{\text{false}\}}{\eta \vdash \{\text{false}\} \cap \{\emptyset\}} \\
(\bot) & \quad \frac{\eta \vdash \{\text{false}\} \cap \{\emptyset\}}{\eta \vdash \{\text{false}\} \cap \{\emptyset\}} \\
(\alpha) & \quad \frac{\eta \vdash \{\alpha \ast \Sigma[j \cup x_j \varphi_q]\} \Sigma[\alpha]\{\phi_i \ast \Sigma[j \cup x_j \varphi_q]\}}{\eta \vdash \{\phi_i \ast \Sigma[j \cup x_j \varphi_q]\} \Sigma[\alpha]\{\text{false}\}} \quad \alpha = (\chi_{p_1,p_2,p_3})
\end{align*}
\]

a. Axioms

\[
\begin{align*}
(1-) & \quad \frac{\eta \backslash \{\Sigma[\alpha]\} \vdash \{\emptyset\} \cap \{\emptyset\}}{\eta \vdash \{\emptyset\} \cap \{\emptyset\}} \\
(1+) & \quad \frac{\eta \cup \{\Sigma[\alpha]\} \vdash \{\emptyset\} \cap \{\emptyset\}}{\eta \vdash \{\emptyset\} \cap \{\emptyset\}} \\
\end{align*}
\]

b. Redundancy Rules

\[
\begin{align*}
(\sigma) & \quad \frac{\eta \vdash \{\phi_i \ast \Sigma \varphi_q\} \Sigma[\alpha]\{\phi_i \ast \Sigma \varphi_q\}}{\eta \vdash \{\phi_i \ast \Sigma \varphi_q\} \Sigma[\alpha]\{\phi_i \ast \Sigma \varphi_q\}} \quad \eta_j = \Sigma[\phi_i \ast \Sigma \varphi_q] \\
\end{align*}
\]

c. Composition Rule

\[
\begin{align*}
(\cdot) & \quad \frac{\eta \vdash \{\phi\} \cap \{\phi\}}{\eta \vdash \{\phi\} \cap \{\phi\}} \\
(\ast) & \quad \frac{\eta \vdash \{\phi\} \cap \{\phi\}}{\eta \vdash \{\phi\} \cap \{\phi\}} \\
\end{align*}
\]

d. Regular Expression Rules

\[
\begin{align*}
(C) & \quad \frac{\eta \vdash \{\phi\} \cap \{\phi'\}}{\eta \vdash \{\phi\} \cap \{\phi'\}} \\
(V) & \quad \frac{\eta \vdash \{\phi\} \cap \{\phi'\} \cap \{\phi\} \cap \{\phi'\} \cap \{\phi\} \cap \{\phi'\}}{\eta \vdash \{\phi\} \cap \{\phi'\} \cap \{\phi\} \cap \{\phi'\} \cap \{\phi\} \cap \{\phi'\}} \\
\end{align*}
\]

e. Structural Rules

Environment assumptions in a rely/assume-guarantee style of reasoning [58, 42]. They are required for soundness, under the semantics of open state changes (Def. 7), which considers that the interactions can fire anytime, unless they are explicitly disabled by some component from \(\phi_i\), for \(i = 1, 2\).
Nevertheless, defining the frontier syntactically faces the following problem: interactions introduced by a predicate atom in $\phi_i$ can impact the state of a component defined by $\phi_{3-i}$. We tackle this problem by forbidding predicate atoms that describe configurations with loose ports, that belong to components lying outside of the current configuration. We recall that a configuration $(C, I, \rho, v)$ is tight if and only if, for each interaction $(c_1, p_1, c_2, p_2) \in I$, we have $c_1, c_2 \in C$. Moreover, we say that a formula $\phi$ is tight if and only if every model of $\phi$ is tight. For instance, a predicate atom $\text{chain}_{h,t}(x,y)$, for given $h,t \geq 0$ (Example 4) is tight, because, in each model, the interactions involve only the $\text{out}$ and $\text{in}$ ports of adjacent components from the configuration.

**Definition 12 (Frontier)** Given symbolic configurations $\phi_1$ and $\phi_2$, the frontier of $\phi_i$ and $\phi_{3-i}$ is the formula $F(\phi_i, \phi_{3-i}) \overset{df}{=} \bigwedge_{\alpha \in \text{inter}(\phi_{3-i}) \setminus (\text{inter}(\phi_i) \cup \text{inter}(\phi_1))} \alpha$, where $\overline{\phi}_i$ is the largest tight subformula of $\phi_i$, for $i = 1,2$.

**Example 8** Let $\phi_1 = \text{chain}_{h,t}(x,y) \text{ out } z \text{ in }$ and $\phi_2 = \text{chain}_{h,t}(y,z) \text{ in } x \text{ out } y$. We have $F(\phi_1, \phi_2) = \langle x, y, z \rangle$ and $F(\phi_2, \phi_1) = \langle y, x, z \rangle$, because the tightness of $\text{chain}_{h,t}(x,y)$ and $\text{chain}_{h,t}(y,z)$ means that the only interactions crossing the boundary of $\phi_1$ and $\phi_2$ are the ones described by $\langle y, x, z \rangle$ and $\langle x, y, z \rangle$.

Finally, the regular expression of the conclusion of the ($\geq$) rule is the interleaving of the regular expressions from the premises, taken with respect to the sets of alphabet symbols $\eta_i = \Sigma[\phi_i \cap F(\phi_i, \phi_{3-i})]$, for $i = 1,2$.

The rules in Fig. 5d introduce regular expressions built using concatenation, Kleene star and union. In particular, for reasons related to the soundness of the proof system, the concatenation rule ($\cdot$) applies to havoc triples whose preconditions are finite disjunctions of symbolic configurations, sharing the same structure of component, interaction and predicate atoms, whereas the cut formulæ (postcondition of the left and precondition of the right premise) share the same structure as the precondition. We formalize below the fact that two formulæ share the same structure:

**Definition 13** Two formulæ $\phi$ and $\psi$ share the same structure, denoted $\phi \simeq \psi$ if and only if they become equivalent when every component atom $x@q$ is replaced by the formula $x@\bot$ in both $\phi$ and $\psi$. We write $\phi \succeq \psi$ if and only if $\phi$ is satisfiable and $\psi$ is not, or else $\phi \simeq \psi$.

The ($\subset$) rule is the dual of ($\cup$), that restricts the language from the conclusion to a subset of the one from the premise. As a remark, by applying the ($\cup$) and ($\subset$) rules in any order, one can derive the havoc invariance of the intermediate assertions in a single-reversal reconfiguration sequence (see Def. 6 and Prop. 1). The rule ($\equiv$) substitutes a regular expression with a congruent one, with respect to the precondition.

Last, the rules in Fig. 5e modify the structure of the pre- and postconditions. In particular, the left unfolding rule (LU) has a premise for each step of unfolding of a predicate atom from the conclusion’s precondition, with respect to a rule from the SID. The environment and the regular expression in each premise are obtained by replacing the alphabet symbol of the unfolded predicate symbol by the set of alphabet symbols from the unfolding step, where $L[\Sigma[\alpha]/L']$ denotes the regular expression obtained by
replacing each occurrence of the alphabet symbol $\Sigma[\alpha]$ in $L$ with the regular expression $L'$.

### 5.3 Havoc Proofs

A **proof tree** is a finite tree $T$ whose nodes are labeled by havoc triples and, for each node $n$ not on the frontier of $T$, the children of $n$ are the premisses of the application of a rule from Fig. 5, whose conclusion is the label of $n$. For the purposes of this paper, we consider only proof trees that meet the following condition:

**Assumption 1** The root of the proof tree is labeled by a havoc triple $\eta \triangleright \{\phi\} L \{\psi\}$, such that $\phi$ is a symbolic configuration and $\eta = \{\Sigma[\alpha] \mid \alpha \in \text{atoms}(\phi)\}$.

It is easy to check that the above condition on the shape of the precondition and the relation between the precondition and the environment holds recursively, for the labels of all nodes in a proof tree that meets assumption 1. Before tackling the soundness of the havoc proof system (Fig. 5), we state an invariance property of the environments of havoc triples that occur in a proof tree:

**Definition 14** A havoc triple $\eta \triangleright \{\phi\} L \{\psi\}$ is **distinctive** if and only if $\langle\langle \Sigma[\alpha_1] \rangle\rangle(\gamma) \cap \langle\langle \Sigma[\alpha_2] \rangle\rangle(\gamma) = \emptyset$, for all $\Sigma[\alpha_1], \Sigma[\alpha_2] \in \eta$ and all $\gamma \in \llbracket\phi\rrbracket$.

The next lemma is proved inductively on the structure of the proof tree, using Assumption 1.

**Lemma 3** Given a proof tree $T$, each node in $T$ is labeled with a distinctive havoc triple.

In order to deal with inductively defined predicates that occur within the pre- and postconditions of the havoc triples, we use cyclic proofs [14]. A **cyclic proof tree** $T$ is a proof tree such that every node on the frontier is either the conclusion of an axiom in Fig. 5a, or there is another node $m$ whose label matches the label of $n$ via a substitution of variables; we say that $n$ is a **bud** and $m$ is its **companion**. A cyclic proof tree is a cyclic proof if and only if every infinite path through the proof tree extended with bud-companion edges, goes through the conclusion of a (LU) rule infinitely often\(^6\). We denote by $\vdash \eta \triangleright \{\phi\} L \{\psi\}$ the fact that $\eta \triangleright \{\phi\} L \{\psi\}$ labels the root of a cyclic proof and state the following soundness theorem:

**Theorem 2** If $\vdash \eta \triangleright \{\phi\} L \{\psi\}$ then $\models \eta \triangleright \{\phi\} L \{\psi\}$.

The proof is by induction on the structure of the proof tree, using Lemma 3.

---

\(^6\)This condition can be effectively decided by checking the emptiness of a Büchi automaton [14].
5.4 A Havoc Proof Example

We demonstrate the use of the proof system in Fig. 5 on the havoc invariance side conditions required by the reconfiguration proofs from §4.3. In fact, we prove a more general statement, namely that chain₁₁(z,x) is havoc invariant, for all h,t ≥ 0. An immediate consequence is that chain₁₁(z,x) is havoc invariant. In particular, the havoc invariance proof for y@H * (y.out, z.in) * chain₁₁(z,x) is an instance of the subgoal (A) below, whereas the proof for y@H * chain₁₁(z,x) can be obtained by applying rules (1+) and (c) to (A), for h = t = 1.

For space reasons, we introduce backlinks from buds to companions whose labels differ by a renaming of free variables and of the h and t indices in chainᵦₜ, such that each pair (h',t') in the label of a companion is lexicographically smaller or equal to a pair (h,t) in the bud. This is a compact (folded) representation of a cyclic proof tree, obtained by repeatedly appending the subtree rooted at the companion to the bud, until all buds are labeled with triples that differ from their companion’s only by a renaming of free variables. Note that such folding is only possible because the definitions of chainᵦᵦ(z,x) and chainᵦᵦᵦ(z,x), for h,t,h',t' ≥ 1 are the same, up to the indices of the predicate symbols (Example 4).

In the proof of the subgoal (A) below, alphabet symbols are abbreviated as Σᵦᵧ ∈ Σ(y.out, y.in) and Σᵦₓᵧ ∈ Σ(chainᵦ₋₁₋₁, y.x). We use the following congruence (Def. 9):

\[(Σᵦₓᵧ ∪ Σᵦₓᵧ) ∗ \equiv z@H*(z.out, y.in) * chainᵦ₋₁₋₁, y.x, Σᵦₓᵧ ∗ \equiv Σᵦₓᵧ ∪ Σᵦₓᵧ * (Σᵦₓᵧ ∪ Σᵦₓᵧ) ∗ \]

The rule (C) strengthens the postcondition chainᵦᵦ(z,x) to an unfolding chainᵦᵦ(z,x) ⊆ ∃ y . z@H *(z.out, y.in) * chainᵦ₋₁₋₁, y.x, whose existentially quantified variable is, moreover, bound to the free variable y from the precondition. The frontier formula in the application of rule (∗) are \(F(z@H, chainᵦ₋₁₋₁, y.x) = F(chainᵦ₋₁₋₁, y.x, z@H) = \text{emp.}\)

\[
(\text{backlink to (I)}) \quad Σᵦₓᵧ ∗ \equiv Σᵦₓᵧ ∪ Σᵦₓᵧ * (Σᵦₓᵧ ∪ Σᵦₓᵧ) ∗
\]

\[
(\text{A1}) \quad Σᵦₓᵧ, Σᵦₓᵧ ∗ \equiv Σᵦₓᵧ ∪ (Σᵦₓᵧ ∪ Σᵦₓᵧ) ∗ \equiv Σᵦₓᵧ ∪ (Σᵦₓᵧ ∪ Σᵦₓᵧ) ∗
\]

\[
(\text{backlink to (I)}) \quad Σᵦₓᵧ, Σᵦₓᵧ ∗ \equiv Σᵦₓᵧ ∪ (Σᵦₓᵧ ∪ Σᵦₓᵧ) ∗ \equiv Σᵦₓᵧ ∪ (Σᵦₓᵧ ∪ Σᵦₓᵧ) ∗
\]

---

This is bound to happen, because a pair (h,t) of positive integers cannot be decreased indefinitely.
right in state leaf
tifiers x state.

The tree architecture guarantees that the notification phase takes time when every component in the system must notify a designated controller, placed in the root of the tree, about an event that involves each component from the frontier of the tree. Conversely, the root component may need to notify the rest of the components.

For space reasons, the proof of the subgoal (B) is provided as supplementary material.

6 A Worked-out Example: Reconfigurable Tree Architectures

In addition to token rings (Fig. 1), we apply our method to reconfiguration scenarios of distributed systems with tree-shaped architectures. Such (virtual) architectures are e.g. used in flooding and leader election algorithms. They are applicable, for instance, when every component in the system must notify a designated controller, placed in the root of the tree, about an event that involves each component from the frontier of the tree. Conversely, the root component may need to notify the rest of the components.

The tree architecture guarantees that the notification phase takes time $O(\log n)$ in the number $n$ of components in the tree, when the tree is balanced, i.e. the lengths of the longest and shortest paths between the root and the frontier differ by at most a constant factor. A reconfiguration of a tree places a designated component (whose priority has increased dynamically) closer to the frontier (dually, closer to the root) in order to receive the notification faster. In balanced trees, reconfigurations involve structure-preserving rotations. For instance, self-adjustable splay-tree networks [62] use the zig (left rotation), zig-zig (left-left rotation) and zig-zag (left-right rotation) operations [65] to move nodes in the tree, while keeping the balance between the shortest and longest paths.

Fig. 6 shows a model of reconfigurable tree architectures, in which each leaf component starts in state leafBusy and sends a notification to its parent before entering the leaf_idle state. An inner component starts in state idle and waits for notifications from both its left ($r_l$) and right ($r_r$) children before sending a notification to its parent ($s$), unless this component is the root (Fig. 6a). We model notifications by interactions of the form $(s, s, w_{r_l})$ and $(s, s, w_{r_r})$. The notification phase is completed when the root is in state right, every inner component is in the idle state and every leaf is in the leaf_Idle state.

Fig. 6b shows a right rotation that reverses the positions of components with identifiers $x$ and $y$, implemented by the reconfiguration program from Fig. 7. The rotation
applies only to configurations in which both \(x\) and \(y\) are in state idle, by distinguishing the case when \(y\) is a left child, the other case being symmetric. Note that, applying the rotation in a configuration where the component indexed by \(x\) is in state right \((\text{both} \ a \ \text{and} \ b \ \text{have sent their notifications to} \ x)\) and the one indexed by \(y\) is in state idle \((c \ \text{has not yet sent its notification to} \ y)\) yields a configuration from which \(c\) cannot send its notification further, because \(x\) has now become the root of the subtree changed by the rotation (a similar scenario is when \(y\) is in state right, \(x\) is in state idle and \(a\), \(b\) and \(c\) have sent their notifications to their parents).

We prove that, whenever a right rotation is applied to a tree, such that the subtrees rooted at \(a\), \(b\) and \(c\) have not sent their notifications yet, the result is another tree in which the subtrees rooted at \(a\), \(b\) and \(c\) are still waiting to submit their notifications. This guarantees that the notification phase will terminate properly with every inner component \((\text{except for the root})\) in state idle and every leaf component in state leafidle, even if one or more reconfigurations take place in between. In particular, this proves the correctness of more complex reconfigurations of splay tree architectures, using e.g. the zig-zig and zig-zag operations [62].

The proof in Fig. 7 uses the inductive definitions from Fig. 6c. The predicates \(\text{tree}_{\text{idle}}(x)\), \(\text{tree}_{\text{idle}}(x)\) define trees where all components are idle, and where some notifications are still being propagated, respectively. The predicate \(\text{tree}(x)\) conveys no
information about the states of the components and the predicate tseg(x,y) defines a tree segment, from component x to component y. To use the havoc proof system from Fig. 5, we need the following statement8:

**Proposition 4** The set of symbolic configurations using predicate atoms tree<sub>idle</sub>(x), tree<sub>−idle</sub>(x), tree(x) and tseg(x,y) is precisely closed.

Moreover, each predicate atom tree<sub>idle</sub>(x), tree<sub>−idle</sub>(x), tree(x) and tseg(x,y) is tight, because, in each model of these atoms, the interactions (u,s,v,r) and (u,s,v,r) are between the ports (s,r) and (s,r) of the components u and v, respectively.

The precondition of the reconfiguration program in Fig. 7 states that x and y are idle components, and the a, b and c subtrees are not idle, whereas the postcondition states

---

8This is similar to Prop. 2.
that the $x$ subtree is not idle. As mentioned, this is sufficient to guarantee the correct termination of the notification phase after the right rotation. As in the proofs from §4.3, proving the correctness of the sequential composition of primitive commands requires proving the havoc invariance of the annotations. However, since in this case, the reconfiguration sequence is single-reversal (Def. 6), we are left with proving havoc invariance only for the annotations marked with $(\sharp)$ in Fig. 7 (Prop. 1). For space reasons, the havoc invariance proofs of these annotations are provided as supplementary material.

7 Towards Automated Proof Generation

Proof generation can be automated, by tackling the following technical problems, briefly described in this section.

The entailment problem  Given a SID $\Delta$ and two CL formulæ $\phi$ and $\psi$, interpreted over $\Delta$, is every model of $\phi$ also a model of $\psi$? This problem arises e.g., when applying the rule of consequence (Fig. 4c bottom-left) in a Hoare-style proof of a reconfiguration program. Unsurprisingly, the CL entailment inherits the positive and negative aspects of the SL entailment [61]. For instance, one can reduce the undecidable problem of universality of context-free languages [5] to CL entailment, with $\phi$ and $\psi$ restricted to predicate atoms. Decidability can be recovered via two restrictions on the syntax of the rules in the SID and a semantic restriction on the configurations that occur as models of the predicate atoms defined by the SID. The syntactic restrictions are that, each rule is of the form $A(x_1, \ldots, x_{\#(A)}) \leftarrow \exists y_1 \ldots \exists y_m \cdot x @ q * \phi * \bigwedge_{\ell=1}^{h} B^f(z_1^{\ell}, \ldots, z_{\#(B)}^{\ell})$, where $\phi$ consists of interaction atoms, such that: (1) $x_1$ occurs in each interaction atom from $\phi$, (2) $\bigcup_{\ell=1}^{h} \{z_1^{\ell}, \ldots, z_{\#(B)}^{\ell}\} = \{x_2, \ldots, x_{\#(A)}\} \cup \{y_1, \ldots, y_m\}$, and (3) for each $\ell \in [1, h]$, $z_1^{\ell}$ occurs in $\phi$. Furthermore, the semantic restriction is that, in each model of a predicate atom, a component must occur in a bounded number of interactions, i.e., the structure is a graph of bounded degree. For instance, star topologies with a central controller and an unbounded number of workers can be defined in CL, but do not satisfy this constraint. With these restrictions, it can be shown that the CL entailment problem is 2EXP-complete, thus matching the complexity of the similar problem for SL [31, 43]. Technical are given in [9].

The frame inference problem  Given two CL formulæ $\phi$ and $\psi$ find a formula $\varphi$, such that $\phi \models \psi \cdot \varphi$. This problem occurs e.g., when applying the frame rule (Fig. 4c bottom-right) with a premiss $\{\phi\} \ R \{\psi\}$ to an arbitrary precondition $\xi$ i.e., one must infer a frame $\varphi$ such that $\xi \models \phi \cdot \varphi$. This problem has been studied for SL [17, 40], in cases where the SID defines only data structures of a restricted form (typically nested lists). Reconsidering the frame inference problem for CL is of paramount importance for automating the generation of Hoare-style correctness proofs and is an open problem.

Automating havoc invariance proofs  Given a precondition $\phi$ and a regular expression $L$, the parallel composition rule ($\bowtie$) requires the inference of regular expressions
L₁ and L₂, such that L₁ ⊲ ⊳ η₁, η₂ L₂ ∼ φ L. We conjecture that, under the bounded degree restriction above, the languages of the frontier (cross-boundary) interactions (Def. 12) are regular and can be automatically inferred by classical automata construction techniques.

8 Related Work

The ability of reconfiguring coordinating architectures of software systems has received much interest in the Software Engineering community, see the surveys [12, 16]. We consider programmed reconfiguration, in which the architecture changes occur according to a sequential program, executed in parallel with the system to which reconfiguration applies. The languages used to write such programs are classified according to the underlying formalism used to define their operational semantics: process algebras, e.g. π-ADL [20], DARWIN [51], hyper-graphs and graph rewriting [66, 69, 50, 19, 3], chemical reactions [68], etc. We separate architectures (structures) from behaviors, thus relating to the BIP framework [6] and its extensions for dynamic reconfigurable systems DR-BIP [32]. In a similar vein, the REO language [4] supports reconfiguration by changing the structure of connectors [24].

Checking the correctness of a dynamically reconfigurable system considers mainly runtime verification methods, i.e. checking a given finite trace of observed configurations against a logical specification. For instance, in [15], configurations are described by annotated hyper-graphs and configuration invariants of finite traces, given first-order logic, are checked using Alloy [41]. More recently, [30, 49, 33] apply temporal logic to runtime verification of reconfigurable systems. Model checking of temporal specifications is also applied to REO programs, under simplifying assumption that render the system finite-state [24]. In contrast, we use induction to deal with parameterized systems of unbounded sizes.

To the best of our knowledge, our work is the first to tackle the verification of reconfiguration programs, by formally proving the absence of bugs, using a Hoare-style annotation of a reconfiguration program with assertions that describe infinite sets of configurations, with unboundedly many components. Traditionally, reasoning about the correctness of unbounded networks of parallel processes uses mostly hard-coded architectures (see [7] for a survey), whereas the more recently developed architecture description logics [44, 52] do not consider the reconfigurability aspect of distributed systems.

Specifying parameterized component-based systems by inductive definitions is not new. Network grammars [64] use context-free grammar rules to describe systems with linear (pipeline, token-ring) structure, obtained by composition of an unbounded number of processes. More complex structures are specified recursively using graph grammar rules with parameter variables [50]. To avoid clashes, these variables must be renamed to unique names and assigned unique indices at each unfolding step. Our recursive specifications use existential quantifiers to avoid name clashes and separating conjunction to guarantee that the components and interactions obtained by the unfolding of the rules are unique.

The assertion language introduced in this paper is a resource logic that supports
local reasoning [56]. Local reasoning about parallel programs has been traditionally within the scope of Concurrent Separation Logic (CSL), that introduced a parallel composition rule [55], with a non-interfering (race-free) semantics of shared-memory parallelism [13]. Considering interference in CSL requires more general proof rules, combining ideas of assume- and rely-guarantee [58, 42] with local reasoning [35, 67] and abstract notions of framing [28, 27, 34]. These rules generalize from both standard CSL parallel composition and rely-guarantee rules, allowing even to reason about properties of concurrent objects, such as (non-)linearizability [63]. However, the body of work on CSL deals almost entirely with shared-memory multithreading programs, instead of distributed systems, which is the aim of our work. In contrast, we develop a resource logic in which the processes do not just share and own resources, but become mutable resources themselves.

9 Conclusions and Future Work

We present a framework for deductive verification of reconfiguration programs, based on a configuration logic that supports local reasoning. We prove the absence of design bugs in ideal networks, without packet loss and communication delays, using a discrete event-based model of behavior, the usual level of abstraction in formal verification of parameterized distributed systems. Our configuration logic relies on inductive predicates to describe systems with unbounded number of components. It is used to annotate reconfiguration programs with Hoare triples, whose validity relies on havoc invariants about the ongoing interactions in the system. These invariants are tackled with a specific proof system, that uses a parallel composition rule in the style of assume/rely-guarantee reasoning.

As future work, we consider push-button techniques for frame inference and havoc invariant synthesis, allowing broadcast interactions between all the components, and extensions of the finite-state model of behavior, using timed and hybrid automata.

References

[1] P. A. Abdulla, G. Delzanno, N. B. Henda, and A. Rezine. Regular model checking without transducers (on efficient verification of parameterized systems). In O. Grumberg and M. Huth, editors, Tools and Algorithms for the Construction and Analysis of Systems, 13th International Conference, TACAS 2007, volume 4424 of LNCS, pages 721–736. Springer, 2007.

[2] C. Aiswarya, B. Bollig, and P. Gastin. An automata-theoretic approach to the verification of distributed algorithms. In L. Aceto and D. de Frutos-Escrig, editors, 26th International Conference on Concurrency Theory, CONCUR 2015, Madrid, Spain, September 1-4, 2015, volume 42 of LIPIcs, pages 340–353. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2015.

[3] C. I. Arad. Programming Model and Protocols for Reconfigurable Distributed Systems. PhD thesis, KTH Royal Institute of Technology, 2013.
[4] F. Arbab. Reo: A channel-based coordination model for component composition. *Mathematical Structures in Comp. Sci.*, 14(3):329–366, June 2004.

[5] Y. Bar-Hillel, M. Perles, and E. Shamir. On formal properties of simple phrase structure grammars. *Sprachtypologie und Universalienforschung*, 14:143–172, 1961.

[6] A. Basu, M. Bozga, and J. Sifakis. Modeling heterogeneous real-time components in BIP. In *Fourth IEEE International Conference on Software Engineering and Formal Methods (SEFM 2006)*, pages 3–12. IEEE Computer Society, 2006.

[7] R. Bloem, S. Jacobs, A. Khalimov, I. Konnov, S. Rubin, H. Veith, and J. Widder. *Decidability of Parameterized Verification*. Synthesis Lectures on Distributed Computing Theory. Morgan & Claypool Publishers, 2015.

[8] B. Bollig, P. Bouyer, and F. Reiter. Identifiers in registers. In M. Bojańczyk and A. Simpson, editors, *Foundations of Software Science and Computation Structures*, pages 115–132, Cham, 2019. Springer International Publishing.

[9] M. Bozga, L. Bueri, and R. Iosif. Decision problems in a logic for reasoning about reconfigurable distributed systems. *CoRR*, abs/2202.09637, 2022.

[10] M. Bozga, J. Esparza, R. Iosif, J. Sifakis, and C. Welzel. Structural invariants for the verification of systems with parameterized architectures. In *Tools and Algorithms for the Construction and Analysis of Systems - 26th International Conference, TACAS 2020*, volume 12078 of LNCS, pages 228–246. Springer, 2020.

[11] M. Bozga and R. Iosif. Specification and safety verification of parametric hierarchical distributed systems. In *Formal Aspects of Component Software - 17th International Conference, FACS 2021, Virtual Event, October 28-29, 2021, Proceedings*, volume 13077 of Lecture Notes in Computer Science, pages 95–114. Springer, 2021.

[12] J. Bradbury, J. Cordy, J. Dingel, and M. Wermelinger. A survey of self-management in dynamic software architecture specifications. In *Proceedings of the 1st ACM SIGSOFT workshop on Self-managed systems*, pages 28–33. ACM, 2004.

[13] S. Brookes and P. W. O’Hearn. Concurrent separation logic. *ACM SIGLOG News*, 3(3):47–65, Aug. 2016.

[14] J. Brotherston and A. Simpson. Sequent calculi for induction and infinite descent. *J. Log. Comput.*, 21(6):1177–1216, 2011.

[15] A. Bucchiarone and J. P. Galeotti. Dynamic software architectures verification using dynalloy. *Electron. Commun. Eur. Assoc. Softw. Sci. Technol.*, 10, 2008.

[16] A. Butting, R. Heim, O. Kautz, J. O. Ringert, B. Rumpe, and A. Wortmann. A classification of dynamic reconfiguration in component and connector architecture description. In *Proceedings of MODELS 2017 Satellite Event: Workshops (ModComp)*, volume 2019 of *CEUR Workshop Proceedings*, pages 10–16. CEUR-WS.org, 2017.
[17] C. Calcagno, D. Distefano, P. W. O’Hearn, and H. Yang. Compositional shape analysis by means of bi-abduction. *J. ACM*, 58(6), Dec. 2011.

[18] C. Calcagno, P. W. O’Hearn, and H. Yang. Local action and abstract separation logic. In 22nd IEEE Symposium on Logic in Computer Science (LICS 2007), 10-12 July 2007, Wroclaw, Poland, Proceedings, pages 366–378. IEEE Computer Society, 2007.

[19] J. Cao, A. T. S. Chan, and Y. Sun. *GOP: A Graph-Oriented Programming Model for Parallel and Distributed Systems*, pages 21–36. Springer US, Boston, MA, 2005.

[20] E. Cavalcante, T. V. Batista, and F. Oquendo. Supporting dynamic software architectures: From architectural description to implementation. In L. Bass, P. Lago, and P. Kruchten, editors, *12th Working IEEE/IFIP Conference on Software Architecture, WICSA 2015*, pages 31–40. IEEE Computer Society, 2015.

[21] E. Chang and R. Roberts. An improved algorithm for decentralized extrema-finding in circular configurations of processes. *Commun. ACM*, 22(5):281–283, May 1979.

[22] Y. Chen, C. Hong, A. W. Lin, and P. Rümmer. Learning to prove safety over parameterised concurrent systems. In D. Stewart and G. Weissenbacher, editors, *2017 Formal Methods in Computer Aided Design, FM CAS 2017*, pages 76–83. IEEE, 2017.

[23] C. Chilton, B. Jonsson, and M. Z. Kwiatkowska. Compositional assume-guarantee reasoning for input/output component theories. *Sci. Comput. Program.*, 91:115–137, 2014.

[24] D. Clarke. A basic logic for reasoning about connector reconfiguration. *Fundam. Inf.*, 82(4):361–390, Feb. 2008.

[25] L. D’Antoni and M. Veanes. Automata modulo theories. *Commun. ACM*, 64(5):86–95, 2021.

[26] S. Demri, É. Lozes, and A. Mansutti. The effects of adding reachability predicates in propositional separation logic. In *FOSSACS 2018, Proceedings*, volume 10803 of *Lecture Notes in Computer Science*, pages 476–493. Springer, 2018.

[27] T. Dinsdale-Young, L. Birkedal, P. Gardner, M. Parkinson, and H. Yang. Views: Compositional reasoning for concurrent programs. *SIGPLAN Not.*, 48(1):287–300, Jan. 2013.

[28] T. Dinsdale-Young, M. Dodds, P. Gardner, M. J. Parkinson, and V. Vafeiadis. Concurrent abstract predicates. In *ECOOP 2010 – Object-Oriented Programming*, pages 504–528, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[29] D. Dolev, M. Klawe, and M. Rodeh. An o(n log n) unidirectional distributed algorithm for extrema finding in a circle. *Journal of Algorithms*, 3(3):245–260, 1982.
[30] J. Dormoy, O. Kouchnarenko, and A. Lanoix. Using temporal logic for dynamic reconfigurations of components. In L. S. Barbosa and M. Lumpe, editors, Formal Aspects of Component Software - 7th International Workshop, FACS 2010, volume 6921 of Lecture Notes in Computer Science, pages 200–217. Springer, 2010.

[31] M. Echenim, R. Iosif, and N. Peltier. Entailment checking in separation logic with inductive definitions is 2-exptime hard. In LPAR 2020: 23rd International Conference on Logic for Programming, Artificial Intelligence and Reasoning, Alicante, Spain, May 22-27, 2020, volume 73 of EPiC Series in Computing, pages 191–211. EasyChair, 2020.

[32] R. El-Ballouli, S. Bensalem, M. Bozga, and J. Sifakis. Programming dynamic reconfigurable systems. International Journal on Software Tools for Technology Transfer, January 2021.

[33] A. El-Hokayem, M. Bozga, and J. Sifakis. A temporal configuration logic for dynamic reconfigurable systems. In C. Hung, J. Hong, A. Bechini, and E. Song, editors, SAC ’21: The 36th ACM/SIGAPP Symposium on Applied Computing, Virtual Event, Republic of Korea, March 22-26, 2021, pages 1419–1428. ACM, 2021.

[34] F. Farka, A. Nanevski, A. Banerjee, G. A. Delbianco, and I. Fábregas. On algebraic abstractions for concurrent separation logics. Proc. ACM Program. Lang., 5(POPL), Jan. 2021.

[35] X. Feng, R. Ferreira, and Z. Shao. On the relationship between concurrent separation logic and assume-guarantee reasoning. In Programming Languages and Systems, pages 173–188. Springer Berlin Heidelberg, 2007.

[36] K. Foerster and S. Schmid. Survey of reconfigurable data center networks: Enablers, algorithms, complexity. SIGACT News, 50(2):62–79, 2019.

[37] I. Foster. What is the grid? a three point checklist. GRID today, 1:32–36, 01 2002.

[38] D. Gall, R. Jacob, A. Richa, C. Scheideler, S. Schmid, and H. Taeubig. A note on the parallel runtime of self-stabilizing graph linearization. Theory of Computing Systems (TOCS), 2014.

[39] D. Gall, R. Jacob, A. Richa, C. Scheideler, S. Schmid, and H. Täubig. A note on the parallel runtime of self-stabilizing graph linearization. Theory of Computing Systems, 55(1):110–135, July 2014.

[40] N. Gorogiannis, M. I. Kanovich, and P. W. O’Hearn. The complexity of abduction for separated heap abstractions. In E. Yahav, editor, Static Analysis - 18th International Symposium, SAS 2011, Venice, Italy, September 14-16, 2011. Proceedings, volume 6887 of Lecture Notes in Computer Science, pages 25–42. Springer, 2011.
[41] D. Jackson. Alloy: a lightweight object modelling notation. *ACM Trans. Softw. Eng. Methodol.*, 11(2):256–290, 2002.

[42] C. B. Jones. *Developing methods for computer programs including a notion of interference*. PhD thesis, University of Oxford, UK, 1981.

[43] J. Katelaan and F. Zuleger. Beyond symbolic heaps: Deciding separation logic with inductive definitions. In *LPAR 2020: 23rd International Conference on Logic for Programming, Artificial Intelligence and Reasoning, Alicante, Spain, May 22-27, 2020*, volume 73 of *EPiC Series in Computing*, pages 390–408. EasyChair, 2020.

[44] I. V. Konnov, T. Kotek, Q. Wang, H. Veith, S. Bliudze, and J. Sifakis. Parameterized systems in BIP: design and model checking. In *27th International Conference on Concurrency Theory, CONCUR 2016*, volume 59 of *LIPIcs*, pages 30:1–30:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016.

[45] C. Krause, Z. Maraikar, A. Lazovik, and F. Arbab. Modeling dynamic reconfigurations in reo using high-level replacement systems. *Sci. Comput. Program.*, 76(1):23–36, 2011.

[46] J. B. Kruskal. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society*, 7(1):48–50, 1956.

[47] L. Lamport. The part-time parliament. *ACM Trans. Comput. Syst.*, 16(2):133–169, may 1998.

[48] L. Lamport, R. Shostak, and M. Pease. The byzantine generals problem. *ACM Trans. Program. Lang. Syst.*, 4(3):382–401, jul 1982.

[49] A. Lanoix, J. Dormoy, and O. Kouchnarenko. Combining proof and model-checking to validate reconfigurable architectures. *Electron. Notes Theor. Comput. Sci.*, 279(2):43–57, 2011.

[50] D. Le Metayer. Describing software architecture styles using graph grammars. *IEEE Transactions on Software Engineering*, 24(7):521–533, 1998.

[51] J. Magee and J. Kramer. Dynamic structure in software architectures. In *ACM SIGSOFT Software Engineering Notes*, volume 21(6), pages 3–14. ACM, 1996.

[52] A. Mavridou, E. Baranov, S. Bliudze, and J. Sifakis. Configuration logics: Modeling architecture styles. *J. Log. Algebr. Meth. Program.*, 86(1):2–29, 2017.

[53] O. Michail, G. Skretas, and P. G. Spirakis. Distributed computation and reconfiguration in actively dynamic networks. In Y. Emek and C. Cachin, editors, *PODC ’20: ACM Symposium on Principles of Distributed Computing, Virtual Event, Italy, August 3-7, 2020*, pages 448–457. ACM, 2020.
[54] M. Noormohammadpour and C. S. Raghavendra. Datacenter traffic control: Understanding techniques and tradeoffs. *IEEE Commun. Surv. Tutorials*, 20(2):1492–1525, 2018.

[55] P. W. O'Hearn. Resources, concurrency, and local reasoning. *Theor. Comput. Sci.*, 375(1-3):271–307, 2007.

[56] P. W. O'Hearn, J. C. Reynolds, and H. Yang. Local reasoning about programs that alter data structures. In *Proceedings of the 15th International Workshop on Computer Science Logic*, CSL ’01, pages 1–19, 2001.

[57] P. W. O’Hearn, H. Yang, and J. C. Reynolds. Separation and information hiding. *ACM Trans. Program. Lang. Syst.*, 31(3), Apr. 2009.

[58] S. Owicki and D. Gries. *An Axiomatic Proof Technique for Parallel Programs*, pages 130–152. Springer New York, New York, NY, 1978.

[59] B. S. Peres, O. A. de Oliveira Souza, O. Goussevskaia, C. Avin, and S. Schmid. Distributed self-adjusting tree networks. In *2019 IEEE Conference on Computer Communications, INFOCOM 2019, Paris, France, April 29 - May 2, 2019*, pages 145–153. IEEE, 2019.

[60] R. C. Prim. Shortest connection networks and some generalizations. *The Bell System Technical Journal*, 36(6):1389–1401, 1957.

[61] J. C. Reynolds. Separation logic: A logic for shared mutable data structures. In *17th IEEE Symposium on Logic in Computer Science (LICS 2002)*, 22-25 July 2002, Copenhagen, Denmark, *Proceedings*, pages 55–74. IEEE Computer Society, 2002.

[62] S. Schmid, C. Avin, C. Scheideler, M. Borokhovich, B. Haeupler, and Z. Lotker. Splaynet: Towards locally self-adjusting networks. *IEEE/ACM Trans. Netw.*, 24(3):1421–1433, June 2016.

[63] I. Sergey, A. Nanevski, A. Banerjee, and G. A. Delbianco. Hoare-style specifications as correctness conditions for non-linearizable concurrent objects. *SIGPLAN Not.*, 51(10):92–110, Oct. 2016.

[64] Z. Shtadler and O. Grumberg. Network grammars, communication behaviors and automatic verification. In J. Sifakis, editor, *Automatic Verification Methods for Finite State Systems, International Workshop*, volume 407 of *LNCS*, pages 151–165. Springer, 1989.

[65] D. D. Sleator and R. E. Tarjan. Self-adjusting binary search trees. *J. ACM*, 32(3):652–686, July 1985.

[66] G. Taentzer, M. Goedicke, and T. Meyer. Dynamic change management by distributed graph transformation: Towards configurable distributed systems. In *International Workshop on Theory and Application of Graph Transformations*, pages 179–193. Springer, 1998.
[67] V. Vafeiadis and M. Parkinson. A marriage of rely/guarantee and separation logic. In CONCUR 2007 – Concurrency Theory, pages 256–271. Springer Berlin Heidelberg, 2007.

[68] M. Wermelinger. Towards a chemical model for software architecture reconfiguration. IEE Proceedings-Software, 145(5):130–136, 1998.

[69] M. Wermelinger and J. L. Fiadeiro. A graph transformation approach to software architecture reconfiguration. Sci. Comput. Program., 44(2):133–155, 2002.

[70] M. Wermelinger, A. Lopes, and J. L. Fiadeiro. A graph based architectural (re)configuration language. SIGSOFT Softw. Eng. Notes, 26(5):21–32, Sept. 2001.
A Proofs from Section 4

Lemma 1 For every program $R \in \mathcal{L}$, the action $\langle\langle R \rangle\rangle$ is local for $\text{modif}(R)$.

Proof. By induction on the structure of the local program $R$. For the base case $R \in \mathcal{Q}$, we check the following points, for all $\gamma_i = (C_i, I_i, \rho_{i \rightarrow v}) \in \Gamma$, for $i = 1, 2$, such that $\gamma_1 \bullet \gamma_2$ is defined:

- $R = \text{new}(q, x)$: we compute $\langle\langle \text{new}(q, x) \rangle\rangle(\gamma_1 \bullet \gamma_2) =$

  $$\langle\langle \text{new}(q, x) \rangle\rangle(\gamma_1 \bullet \gamma_2) = \langle\langle C_1 \cup C_2 \cup \{c\}, I_1 \cup I_2, (\rho_1 \cup \rho_2)[c \leftarrow q, v[x \leftarrow c]\rangle\rangle = \langle\langle C_1 \cup \{c\}, I_1, \rho_1[c \leftarrow q, v[x \leftarrow c]\rangle\rangle \cup \langle\langle C_2, I_2, \rho_2, v[x \leftarrow c]\rangle\rangle \subseteq \langle\langle \text{new}(q, x) \rangle\rangle(\gamma_1) \bullet \{\gamma_2\} \uparrow^{\text{modif}(\text{new}(q, x))}$$

- $R = \text{delete}(x)$: we distinguish the following cases:
  - if $v(x) \in C_1$, we compute $\langle\langle \text{delete}(x) \rangle\rangle(\gamma_1 \bullet \gamma_2) =$

    $$\langle\langle \text{delete}(x) \rangle\rangle(\gamma_1 \bullet \gamma_2) = \langle\langle C_1 \setminus \{v(x)\}, I_1 \cup I_2, (\rho_1 \cup \rho_2, v)\rangle\rangle \cup \langle\langle C_2, I_2, \rho_2, v[x \leftarrow c]\rangle\rangle \subseteq \langle\langle \text{delete}(x) \rangle\rangle(\gamma_1) \bullet \{\gamma_2\} \uparrow^{\text{modif}(\text{delete}(x))}$$

  - else $v(x) \notin C_1$ and $\langle\langle \text{delete}(x) \rangle\rangle(\gamma_1) = \top$, thus we obtain:

    $$\langle\langle \text{delete}(x) \rangle\rangle(\gamma_1 \bullet \gamma_2) \subseteq \top \bullet \{\gamma_2\} \uparrow^{\text{modif}(\text{delete}(x))} = \langle\langle \text{delete}(x) \rangle\rangle(\gamma_1) \bullet \{\gamma_2\} \uparrow^{\text{modif}(\text{delete}(x))}$$

- $R = \text{connect}(x_1, p_1, x_2, p_2)$: we compute $\langle\langle \text{connect}(x_1, p_1, x_2, p_2) \rangle\rangle(\gamma_1 \bullet \gamma_2) =$

  $$\langle\langle \text{connect}(x_1, p_1, x_2, p_2) \rangle\rangle(\gamma_1 \bullet \gamma_2) = \langle\langle C_1 \cup C_2 \cup \{v(x_1), p_1, v(x_2), p_2\}, I_1 \cup I_2, (\rho_1 \cup \rho_2, v)\rangle\rangle \cup \langle\langle C_2, I_2, \rho_2, p_1 \cup p_2, v\rangle\rangle \subseteq \langle\langle \text{connect}(x_1, p_1, x_2, p_2) \rangle\rangle(\gamma_1) \bullet \{\gamma_2\} \uparrow^{\text{modif}(\text{connect}(x_1, p_1, x_2, p_2))}$$

- $R = \text{disconnect}(x_1, p_1, x_2, p_2)$: we distinguish the following cases:
  - if $\langle\langle \text{disconnect}(x_1, p_1, x_2, p_2) \rangle\rangle(\gamma_1 \bullet \gamma_2) =$

    $$\langle\langle \text{disconnect}(x_1, p_1, x_2, p_2) \rangle\rangle(\gamma_1 \bullet \gamma_2) = \langle\langle C_1 \cup C_2 \cup \{v(x_1), p_1, v(x_2), p_2\}, \rho_1 \cup \rho_2, v\rangle\rangle \cup \langle\langle C_2, I_2, \rho_2, p_1 \cup p_2, v\rangle\rangle \subseteq \langle\langle \text{disconnect}(x_1, p_1, x_2, p_2) \rangle\rangle(\gamma_1) \bullet \{\gamma_2\} \uparrow^{\text{modif}(\text{disconnect}(x_1, p_1, x_2, p_2))}$$

  - else $\langle\langle \text{disconnect}(x_1, p_1, x_2, p_2) \rangle\rangle(\gamma_1 \bullet \gamma_2) = \top$, thus:

    $$\langle\langle \text{disconnect}(x_1, p_1, x_2, p_2) \rangle\rangle(\gamma_1 \bullet \gamma_2) \subseteq \top \bullet \{\gamma_2\} = \langle\langle \text{disconnect}(x_1, p_1, x_2, p_2) \rangle\rangle(\gamma_1) \bullet \{\gamma_2\} \uparrow^{\text{modif}(\text{disconnect}(x_1, p_1, x_2, p_2))}$$

- $R = \text{skip}$: this case is a trivial check.
For the inductive step, we check the following points:

- $R = R_1 + R_2$: we compute $\langle\langle R_1 + R_2 \rangle\rangle(\gamma_1 \cdot \gamma_2) =$

  $\langle\langle R_1 \rangle\rangle(\gamma_1 \cdot \gamma_2) \cup \langle\langle R_2 \rangle\rangle(\gamma_1 \cdot \gamma_2) \subseteq \{\text{by the inductive hypothesis}\}$

  $\langle\langle R_1 \rangle\rangle(\gamma_1) \cdot \{\gamma_2\}^{\text{modif}(R_1)} \cup \langle\langle R_2 \rangle\rangle(\gamma_1) \cdot \{\gamma_2\}^{\text{modif}(R_2)} \subseteq \text{modif}(R_1 + R_2) = \text{modif}(R_1) \cup \text{modif}(R_2)$

  $\langle\langle R_1 \rangle\rangle(\gamma_1) \cdot \{\gamma_2\}^{\text{modif}(R_1 + R_2)} \cup \langle\langle R_2 \rangle\rangle(\gamma_1) \cdot \{\gamma_2\}^{\text{modif}(R_1 + R_2)} =

  \langle\langle R_1 \rangle\rangle(\gamma_1) \cup \langle\langle R_2 \rangle\rangle(\gamma_1) \cdot \{\gamma_2\}^{\text{modif}(R_1 + R_2)}$

- $R = (\text{with } x : \pi \text{ do } R_1 \text{ od})$, where $\pi$ consists of equalities and disequalities: we distinguish the cases below:

  - if $\gamma_1 \cdot \gamma_2 \models \pi$, we compute

    $\langle\langle \text{with } x : \pi \text{ do } R_1 \text{ od} \rangle\rangle(\gamma_1 \cdot \gamma_2) \subseteq \langle\langle R_1 \rangle\rangle(\gamma_1 \cdot \gamma_2) = \langle\langle R_1 \rangle\rangle(\gamma_1) \cdot \{\gamma_2\}^{\text{modif}(\text{with } x : \pi \text{ do } R_1 \text{ od})}$

  - else $\gamma_1 \cdot \gamma_2 \not\models \pi$ and

    $\langle\langle \text{with } x : \pi \text{ do } R_1 \text{ od} \rangle\rangle(\gamma_1 \cdot \gamma_2) = \emptyset \subseteq \langle\langle R_1 \rangle\rangle(\gamma_1) \cdot \{\gamma_2\}^{\text{modif}(\text{when } x \pi \text{ do } R_1)}$

$\square$

**Lemma 2** For each axiom $\{\phi\} R \{\psi\}$, where $R \in \mathcal{P}$ is primitive, we have $\langle\langle R \rangle\rangle(\langle\langle\phi\rangle\rangle) = \langle\langle\psi\rangle\rangle$.

**Proof.** The proof goes by case split on the type of the primitive command $R$, which determines the pre- and post-condition $\phi$ and $\psi$ of the axiom, respectively:

- $R = \text{new}(q, x)$, $\phi = \text{emp}$ and $\psi = x@q$:

  $\langle\langle \text{new}(q, x) \rangle\rangle(\langle\langle\text{emp}\rangle\rangle) = \{\{c\}, \emptyset, p[c \gets q] \mid c \in \mathbb{C}\} = \langle\langle x@q \rangle\rangle$

  The second step applies the definition $\langle\langle R \rangle\rangle(\langle\langle\phi\rangle\rangle) = \{\gamma \mid \exists \gamma \in \langle\langle\phi\rangle\rangle.R : \gamma \rightarrow \gamma\}$ to the case $R = \text{new}(q, x)$, where the judgement $\text{new}(q, x) : \gamma \rightarrow \gamma'$ is defined in Fig. 3. The rest is by the semantics of CL.

- $R = \text{delete}(x)$, $\phi = x@_-$ and $\psi = \text{emp}$:

  $\langle\langle \text{delete}(x) \rangle\rangle(\langle\langle x@_- \rangle\rangle) = \langle\langle \text{delete}(x) \rangle\rangle(\langle\langle \{v(x)\}, \emptyset, p, v \mid \text{dom}(p) = \{v(x)\}\rangle\rangle) = \langle\langle\text{emp}\rangle\rangle$

  The second step applies the definition $\langle\langle R \rangle\rangle(\langle\langle\phi\rangle\rangle) = \{\gamma \mid \exists \gamma \in \langle\langle\phi\rangle\rangle.R : \gamma \rightarrow \gamma\}$ to the case $R = \text{delete}(x)$, where the judgement $\text{delete}(x) : \gamma \rightarrow \gamma'$ is defined in Fig. 3. The rest is by the semantics of CL.

- $R = \text{connect}(x_1, p_1, x_2, p_2)$, $\phi = \text{emp}$ and $\psi = \langle x_1, p_1, x_2, p_2 \rangle$: similar to $R = \text{new}(q, x)$.
• \( R = \text{disconnect}(x_1.p_1, x_2.p_2), \phi = \langle x_1.p_1, x_2.p_2 \rangle \) and \( \psi = \text{emp} \): similar to \( R = \text{delete}(x) \).

• \( R = \text{skip} \) and \( \phi = \psi = \text{emp} \): trivial. □

**Theorem 1** For any Hoare triple \( \{ \phi \} R \{ \psi \} \), if \( \vdash \{ \phi \} R \{ \psi \} \) then \( \models \{ \phi \} R \{ \psi \} \).

**Proof.** We prove that the inference rules in Fig. 4 are sound. For the axioms, soundness follows from Lemma 2. The rules for the composite programs are proved below by a case split on the syntax of the program from the conclusion \( \{ \phi \} R \{ \psi \} \), assuming that \( \models \{ \phi \} R_i \{ \psi_i \} \), for each premiss \( \{ \phi_i \} R_i \{ \psi_i \} \) of the rule:

• \( R = \text{with} x_1, \ldots, x_k :: \phi \) do \( R \) od: let \( (C, I, \rho, \nu) \in \llbracket R \rrbracket \) be a configuration and distinguish the following cases:

  - if \( (C, I, \rho, \nu[x_1 <- c_1, \ldots, x_k <- c_k]) \models \phi \land \text{true}, \) for some \( c_1, \ldots, c_k \in C \), we obtain \( (C, I, \rho, \nu[x_1 <- c_1, \ldots, x_k <- c_k]) \models \phi \land \text{true} \), because \( \text{fv}(\phi) \cap \{x_1, \ldots, x_k\} = \emptyset \). Then \( (C, I, \rho, \nu \models (C, I, \rho, \nu[x_1 <- c_1, \ldots, x_k <- c_k]) \subseteq \llbracket \psi \rrbracket \supseteq \llbracket \exists x. \psi \rrbracket \) follows from the premiss of the rule.

  - otherwise, we have \( (C, I, \rho, \nu) \models \forall x_1 \ldots \forall x_k. \neg(\phi \land \text{true}) \) and \( (C, I, \rho, \nu) \models \forall x. \exists \nu. \phi \) follows.

• the cases \( R = R_1; R_2, R = R_1 + R_2 \) and \( R = R_1^+ \) are simple checks using the operational semantics rules from Fig. 3.

Concerning the structural rules, we show only the soundness of the frame rule below; the other rules are simple checks, left to the reader. Let \( \gamma \in \llbracket \phi \land \psi \rrbracket \) be a configuration. By the semantics of \( * \), there exists \( \gamma_1 \in \llbracket \phi \rrbracket \) and \( \gamma_2 \in \llbracket \psi \rrbracket \), such that \( \gamma = \gamma_1 \bullet \gamma_2 \). Since \( R \in 2 \), by Lemma 1, we obtain \( \langle R \rangle (\gamma_1 \bullet \gamma_2) \subseteq \langle R \rangle (\gamma_1) \bullet \text{modif}(R) \). Since \( \gamma_1 \in \llbracket \phi \rrbracket \), by the hypothesis on the premiss we obtain \( \langle R \rangle (\gamma_1) \subseteq \llbracket \phi \rrbracket \). Moreover, since \( \gamma_2 \in \llbracket \psi \rrbracket \) and \( \text{modif}(R) \cap \text{fv}(\phi) \), we obtain \( \gamma_2 \bullet \text{modif}(R) \subseteq \llbracket \psi \rrbracket \), leading to \( \langle R \rangle (\gamma) \subseteq \llbracket \psi \land \phi \rrbracket \), as required. □

**Proposition 1** Let \( R = \text{disconnect}(x_1.p_1, x'_1.p'_1); \ldots; \text{disconnect}(x_k.p_k, x'_k.p'_k); \text{connect}(x_{k+1}.p_{k+1}, x'_{k+1}.p'_{k+1}); \ldots; \text{connect}(x_{\ell}.p_{\ell}, x'_{\ell}.p'_{\ell}) \) be a single reversal program. If \( \phi_0, \ldots, \phi_\ell \) are CL formulæ, such that:

• \( \models \{ \phi_{i-1} \} \text{disconnect}(x_i.p_i, x'_i.p'_i) \{ \phi_i \}, \) for all \( i \in [1, k] \),

• \( \models \{ \phi_{j-1} \} \text{connect}(x_j.p_j, x'_j.p'_j) \{ \phi_j \}, \) for all \( j \in [k+1, \ell] \), and

• \( \phi_1 \) and \( \phi_{\ell-1} \) are havoc invariant,

then we have \( \models \{ \phi_0 \} R \{ \phi_\ell \} \).

**Proof.** In order to apply the sequential composition rule for the entire sequence, we need to prove that \( \phi_1, \ldots, \phi_{\ell-1} \) are havoc invariant:
• For $i \in [1, k]$, the proof is by induction on $i$. In the base case, $\phi_1$ is havoc invariant, by the hypothesis. For the inductive step $i \in [2, k]$, let $(C, I, \rho, v) \models \phi_i$ and $(C, I, \rho, v) \Rightarrow \cdots \Rightarrow (C, I, \rho', v)$ be a sequence of state changes induced by the execution of some interactions $(c_{1i}, p_{1i}, c'_{1i}, p'_{1i}), \ldots, (c_{ni}, p_{ni}, c'_{ni}, p'_{ni}) \in I$. Since $\models \{ \phi_{i-1} \}$ disconnect $(x_i, p_i, x'_i, p'_i) \{ \phi_i \}$, by the hypothesis, there exists a model $(C, I', \rho, v)$ of $\phi_{i-1}$, such that $(c_{1i}, p_{1i}, c'_{1i}, p'_{1i}), \ldots, (c_{ni}, p_{ni}, c'_{ni}, p'_{ni}) \in I'$. Since $\phi_{i-1}$ is havoc invariant, by the inductive hypothesis, we have $(C, I', \rho', v) \models \phi_{i-1}$ and, since $I = I' \setminus \{ (v(x_i), p_i, v(x'_i), p'_i) \}$, we have $(v(x_i), p_i, v(x'_i), p'_i) \notin \{ (c_{1i}, p_{1i}, c'_{1i}, p'_{1i}), \ldots, (c_{ni}, p_{ni}, c'_{ni}, p'_{ni}) \}$, thus $(C, I, \rho', v) \models \phi_i$. Since the choices of $(C, I, \rho, v)$ and $(c_{1i}, p_{1i}, c'_{1i}, p'_{1i}), \ldots, (c_{ni}, p_{ni}, c'_{ni}, p'_{ni})$ were arbitrary, we obtain that $\phi_i$ is havoc invariant.

• For $i \in [k + 1, \ell - 1]$, the proof is by reversed induction on $i$. In the base case, $\phi_{\ell-1}$ is havoc invariant, by the hypothesis. For the inductive step, let $(C, I, \rho, v) \models \phi_{\ell-1}$ and $(C, I, \rho, v) \Rightarrow \cdots \Rightarrow (C, I, \rho', v)$ be a sequence of state changes, induced by the executions of some interactions $(c_{1i}, p_{1i}, c'_{1i}, p'_{1i}), \ldots, (c_{ni}, p_{ni}, c'_{ni}, p'_{ni}) \in I$. Since $\models \{ \phi_{\ell-1} \}$ connect $(x_i, p_i, x'_i, p'_i) \{ \phi_i \}$, by the hypothesis, there exists a model $(C, I', \rho, v)$ of $\phi_i$, such that $(c_{1i}, p_{1i}, c'_{1i}, p'_{1i}), \ldots, (c_{ni}, p_{ni}, c'_{ni}, p'_{ni}) \in I'$. Since $\phi_i$ is havoc invariant, by the inductive hypothesis, we have $(C, I', \rho', v) \models \phi_i$ and, since $I = I' \setminus \{ (v(x_i), p_i, v(x'_i), p'_i) \}$, we have $(v(x_i), p_i, v(x'_i), p'_i) \notin \{ (c_{1i}, p_{1i}, c'_{1i}, p'_{1i}), \ldots, (c_{ni}, p_{ni}, c'_{ni}, p'_{ni}) \}$ and $(C, I, \rho', v) \models \phi_{\ell-1}$. Since the choices of $(C, I, \rho, v)$ and $(c_{1i}, p_{1i}, c'_{1i}, p'_{1i}), \ldots, (c_{ni}, p_{ni}, c'_{ni}, p'_{ni})$ were arbitrary, $\phi_{\ell-1}$ is havoc invariant.

\[ \square \]

B Proof from Section 5

**Proposition 2** The set of symbolic configurations built using predicate atoms chain_{h,t}(x, y), for $h, t \geq 0$ (Example 4) is precisely closed.

**Proof.** Let $\phi_i \equiv \phi \ast \bigwedge_{j=1}^{k_i} \text{chain}_{h_{i,j}, t_{i,j}}(x_{i,j}, y_{i,j})$ be symbolic configurations, where $\phi$ is a predicate-free symbolic configuration and $h_{i,j}, t_{i,j} \geq 0$ are integers, for all $j \in [1, k_i]$ and $i = 1, 2$. We prove that $\phi_1$ is precise on $\|\phi_2\|$. Let $\gamma = (C, I, \rho, v) \in \|\phi_2\|$ be a configuration and suppose that there exist configurations $\gamma' = (C', I', \rho', v)$ and $\gamma'' = (C'', I'', \rho'', v)$, such that $\gamma' \subseteq \gamma$, $\gamma'' \subseteq \gamma$, $\gamma' \models \phi_1$ and $\gamma'' \models \phi_1$. Then there exist configurations $\gamma'_0 \equiv (C'_0, I'_0, \rho, v), \ldots, \gamma'_{k_i} \equiv (C'_{k_i}, I'_{k_i}, \rho, v)$ and $\gamma''_0 \equiv (C''_0, I''_0, \rho, v), \ldots, \gamma''_{k_i} \equiv (C''_{k_i}, I''_{k_i}, \rho, v)$, such that:

\[ \gamma' = \bigoplus_{j=0}^{k_i} \gamma'_j \text{ and } \gamma'' = \bigoplus_{j=0}^{k_i} \gamma''_j, \]

\[ \gamma'_0 \models \phi_1 \text{ and } \gamma''_0 \models \phi_1, \]

\[ \gamma'_j \models \text{chain}_{h_{1,j}, t_{1,j}}(x_{1,j}, y_{1,j}) \text{ and } \gamma''_j \models \text{chain}_{h_{1,j}, t_{1,j}}(x_{1,j}, y_{1,j}), \text{ for all } j \in [1, k_i]. \]

Since $\phi_1$ is a predicate-free symbolic configuration, we have $C'_0 = C''_0$ and $I'_0 = I''_0$, thus $\gamma'_0 = \gamma''_0$. Moreover, for each $j \in [1, k_i]$, we have $C'_j = C''_j$ and $I'_j = I''_j$, because both configurations consist of the tight interactions $(c_1, out, c_2, in), \ldots, (c_{\ell-1}, out, c_\ell, in)$, such
that \( v(x_{1,j}) = c_1 \) and \( v(y_{1,j}) = c_\ell \). Thus, we obtain \( \gamma_j' = \gamma_j'' \), for all \( j \in [0,k] \), leading to \( \gamma' = \gamma'' \). □

**Proposition 3** If \( \models \eta \triangleright \{ \phi \} \Sigma[\phi]^* \{ \psi \} \) then \( h(\{ \phi \}) \subseteq \{ \psi \} \).

**Proof.** Let \( \gamma = (C, I, \rho, v) \) be a model of \( \phi \) i.e., \( \gamma \in \{ \phi \} \). It is sufficient to prove that \( h(\gamma) \subseteq \{ (C, I, \rho', v) | \gamma \triangleright w \rightarrow (C, I, \rho', v), w \in \{ \Sigma[\phi]^* \} \} \), because \( \{ (C, I, \rho', v) | \rho \triangleright w \rightarrow (C, I, \rho', v) \} \subseteq \{ \psi \} \) for each \( w \in \{ \Sigma[\phi]^* \} \), by the hypothesis \( \models \eta \triangleright \{ \phi \} \Sigma[\phi]^* \{ \psi \} \) (Def. 10).

Let \( \gamma' = (C, I, \rho', v) \in h(\gamma) \) be a configuration. Then there exists a finite sequence of interactions, say \( w = (c_1, p_1, c_1', p_1') \ldots (c_n, p_n, c_n', p_n') \in I^* \), such that \( (C, I, \rho, v) \triangleright \rightarrow (C, I, \rho', v) \), by Def. 3. Note that \( \{ (C, I, \rho'', v) | \gamma' \triangleright \rightarrow (C, I, \rho'', v) \} = \{ (C, I, \rho, v) \triangleright \rightarrow (C, I, \rho'', v) \} \subseteq \{ (C, I, \rho', v) | \rho \triangleright w \rightarrow (C, I, \rho', v), w \in \{ \Sigma[\phi]^* \} \} \), for each interaction \( (c_1, p_1, c_2, p_2) \in I \), by Def. 3 and Def. 7. It remains to show that \( w \in \{ \Sigma[\phi]^* \} \). Since \( (C, I, \rho, v) \models \phi \), for any interaction \( (c_k, p_k, c_k', p_k') \in I \), for \( k \in [1,n] \), we distinguish two cases, either:

- there exists an interaction atom \( \alpha = (x_k, p_k, c_k', p_k') \in \text{inter}(\phi) \) and a configuration \( (C'', I'', \rho'', v) \subseteq \gamma \), such that \( I'' = \{ (c_k, p_k, c_k', p_k') \} \) and \( (C'', I'', \rho'', v) \models \alpha \), or
- there exists a predicate atom \( \alpha \in \text{pred}(\phi) \) and a configuration \( (C'', I'', \rho'', v) \subseteq \gamma \), such that \( (c_k, p_k, c_k', p_k') \in I'' \) and \( (C'', I'', \rho'', v) \models \alpha \).

In both cases, we have \( (c_k, p_k, c_k', p_k') \in \{ \Sigma[\phi] \}(\gamma) \), for some \( \alpha \in \text{atoms}(\phi) \), thus \( (c_k, p_k, c_k', p_k') \in \{ \Sigma[\phi] \}(\gamma) \), because \( \Sigma[\phi] = \bigcup_{\alpha \in \text{atoms}(\phi)} \Sigma[\alpha] \). Since the choice of \( k \in [1,n] \) is arbitrary, we obtain that \( w \in \{ \Sigma[\phi]^* \}(\gamma) \). □

**Lemma 3** Given a proof tree \( T \), each node in \( T \) is labeled with a distinctive havoc triple.

**Proof.** The proof goes by induction on the structure of the proof tree. For the base case, the tree consists of a single root node and let \( \eta \triangleright \{ \phi \} L \{ \psi \} \) be the label of the root node. By Assumption 1, \( \phi \) is a symbolic configuration and \( \eta = \{ \Sigma[\alpha_1], \ldots, \Sigma[\alpha_k] \} \), where \( \text{atoms}(\phi) = \{ \alpha_1, \ldots, \alpha_k \} \) is the set of interaction and predicate atoms from \( \phi \).

Let \( \gamma \) be a model of \( \phi \), hence there exist configurations \( \gamma_0, \gamma_1, \ldots, \gamma_k \), such that \( \gamma = \bigotimes_{i=0}^k \gamma_i \) and \( \gamma_i \models \alpha_i \), for all \( i \in [1,k] \). Because the composition \( \bigotimes \), is defined, we obtain that \( \{ \Sigma[\alpha_i] \}(\gamma) \cap \{ \Sigma[\alpha_j] \}(\gamma) = \emptyset \), for all \( i \neq j \in [1,k] \). Moreover, since each formula \( \alpha_i \in \text{atoms}(\phi) \) is precise on \( \{ \phi \} \), we have \( \{ \Sigma[\alpha_i] \}(\gamma_i) = \{ \Sigma[\alpha_i] \}(\gamma) \), hence \( \{ \Sigma[\alpha_i] \}(\gamma) \cap \{ \Sigma[\alpha_j] \}(\gamma) = \emptyset \), for all \( i \neq j \in [1,k] \). For the inductive step, we distinguish the cases below, based on the type of the inference rule that expands the root:

- (1) Let \( \eta \triangleright \{ \phi \}, \{ x_1, p_1, x_2, p_2 \} \) be label of the root of the proof tree and \( \gamma \in \{ \phi \}, \{ x_1, p_1, x_2, p_2 \} \) be a configuration. Then there exists configurations \( \gamma_0 \) and \( \gamma_1 \), such that \( \gamma = \gamma_0 \cdot \gamma_1 \) and \( \gamma_0 \models \phi \) and \( \gamma_1 \models (x_1, p_1, x_2, p_2) \). By Assumption 1, we have \( \eta = \Sigma[\phi] = \{ \Sigma[x_1, p_1, x_2, p_2] \} \). By the inductive hypothesis, the premiss \( \eta \triangleright \{ x_1, p_1, x_2, p_2 \} \triangleright \{ \phi \} \) of the rule is distinctive, hence the interpretations of the atoms in the environment \( \{ \Sigma[\alpha_i] \}(\gamma_0) \) are pairwise disjoint. Since each predicate atom \( \alpha \in \text{atoms}(\phi) \) is precise on \( \{ \phi \} \),
the sets \( \langle \Sigma[A]\rangle(\gamma), \alpha \in \text{atoms}(\psi) \) are also pairwise disjoint. Since \( \langle x_1, p_1, x_2, p_2 \rangle \) is precise on \( \Gamma \), we obtain that \( \langle \{ x_1, p_1, x_2, p_2 \}\rangle(\gamma_1) = \langle \{ x_1, p_1, x_2, p_2 \}\rangle(\gamma) \) and, since \( \gamma = \gamma_0 \cdot \gamma_1 \), the set \( \langle \{ x_1, p_1, x_2, p_2 \}\rangle(\gamma) \) is disjoint from the sets \( \langle \Sigma[A]\rangle(\alpha), \alpha \in \text{atoms}(\phi) \), thus \( \eta \bowtie \{ \phi * \{ x_1, p_1, x_2, p_2 \}\} L \{ \psi * \{ x_1, p_1, x_2, p_2 \}\} \) is distinctive.

• (L+) Let \( \eta \bowtie \{ \phi \} L \{ \psi \} \) be the root label and let \( \gamma \in \{ \phi \} \) be a configuration. By Assumption 1, we have \( \eta = \Sigma[\phi] \) and let \( \{ x_1, p_1, x_2, p_2 \} \) be an interaction atom, such that \( \phi \bowtie \{ x_1, p_1, x_2, p_2 \} \). Let \( \gamma' \) be any model of \( \{ x_1, p_1, x_2, p_2 \} \). By \( \phi \bowtie \{ x_1, p_1, x_2, p_2 \} \), we have \( \{ x_1, p_1, x_2, p_2 \} \notin \eta \) and, moreover, the composition \( \gamma \cdot \gamma' \) is defined, thus \( \gamma \cdot \gamma' \in \{ \phi \cdot \{ x_1, p_1, x_2, p_2 \}\} \). By the inductive hypothesis, \( \eta \cup \{ \Sigma[\{ x_1, p_1, x_2, p_2 \}] \} \bowtie \{ \phi \cdot \{ x_1, p_1, x_2, p_2 \}\} L \{ \psi * \{ x_1, p_1, x_2, p_2 \}\} \) is distinctive, hence \( \{ \Sigma[\alpha] \}(\gamma \cdot \gamma') \cap \{ \Sigma[\alpha] \}(\gamma \cdot \gamma') = \emptyset \), for all \( \Sigma[\alpha] \), \( \Sigma[\alpha] \in \eta \). Since \( \gamma' \bowtie \{ x_1, p_1, x_2, p_2 \} \), \( \eta = \Sigma[\phi] \) and \( \phi \bowtie \{ x_1, p_1, x_2, p_2 \} \), we obtain \( \{ \Sigma[\alpha] \}(\gamma \cdot \gamma') \) for all \( \Sigma[\alpha] \in \eta \), thus \( \eta \bowtie \{ \phi \} L \{ \psi \} \) is distinctive.

• (oo) Let \( \eta_1 \cup \eta_2 \bowtie \{ \phi_1 \cdot \phi_2 \} L_1 \bowtie \eta_1, \eta_2 L_2 \bowtie \{ \psi_1 \cdot \psi_2 \} \) be the label of the root, \( \eta_1 = \Sigma[\phi_1] \cup \{ \phi \} \), for \( i = 1, 2 \), and let \( \gamma \) be a model of the preconditional of this havoc triple. Then there exist two configurations \( \gamma_1, \gamma_2 \), such that \( \gamma_1 \cdot \gamma_2 \) and \( \gamma_1 \bowtie \phi_1 \), for \( i = 1, 2 \). By Assumption 1, we have \( \eta_1 \cup \eta_2 = \Sigma[\phi_1] \cup \Sigma[\phi_2] \). Let \( \gamma' \) be a structure, such that \( \gamma' \subseteq \gamma_1 \) and \( \gamma' \bowtie \{ \phi \} \), for \( i = 1, 2 \). By the definition of \( \{ \phi \} \), as separated conjunction of interaction atoms from \( \phi_{3-} \), these substructures exist, and moreover, because each interaction atom is precise on \( \gamma \), they are unique. Then we have \( \gamma' \bowtie \phi \bowtie \{ \phi \} \), for \( i = 1, 2 \). By the inductive hypothesis, since each havoc triple \( \eta_1 \bowtie \{ \phi \cdot \{ \phi, \phi_{3-} \} \} L_1 \bowtie \{ \psi_1 \cdot \{ \psi, \phi_{3-} \} \} \) is distinctive, the sets \( \{ \Sigma[\alpha] \}(\gamma \cdot \gamma') \) for \( \alpha \in \text{atoms}(\phi) \) are pairwise disjoint, for \( i = 1, 2 \). Since each predicate atom \( \alpha \) \( \in \text{atoms}(\phi \cdot \phi_2) \) is precise on \( \{ \phi \cdot \phi_2 \} \), hence the sets \( \{ \Sigma[\alpha] \}(\gamma \cdot \gamma') \) are pairwise disjoint as well, for \( i = 1, 2 \). Since the configurations \( \gamma_1, \gamma_2 \) share no interactions, the havoc triple \( \eta_1 \cup \eta_2 \bowtie \{ \phi_1 \cdot \phi_2 \} L_1 \bowtie \eta_1, \eta_2 L_2 \bowtie \{ \psi_1 \cdot \psi_2 \} \) is distinctive.

• (LU) Let \( \eta \bowtie \{ \phi \cdot A(y_1, \ldots, y_{n(A)}) \} L \{ \psi \} \) be the label of the root let \( \gamma \) be a model of the preconditional of this havoc triple. Then there exist configurations \( \gamma_0 = (C_0, I_0, I_0, \psi, \Gamma) \) and \( \gamma_1 = (C_1, I_1, I_1, \psi, \Gamma) \), such that \( \gamma = \gamma_0 \cdot \gamma_1 \bowtie \phi \) and \( \gamma_1 = A(y_1, \ldots, y_{n(A)}) \). By Assumption 1, we have \( \eta = \Sigma[\phi] \cup \{ \Sigma[A(y_1, \ldots, y_{n(A)})] \} \). Since \( \gamma_1 = A(y_1, \ldots, y_{n(A)}) \), there exists a rule \( A(x_1, \ldots, x_{n(A)}) \leftarrow \exists z_1 \ldots \exists z_{n(A)} \cdot \phi \) in the SID, where \( \phi \) is a symbolic configuration, such that \( (C_1, I_1, I_1, \psi, \Gamma) \bowtie \{ \psi \} \) for some components \( c_1, \ldots, c_{n(A)} \in \mathbb{C} \), and let \( \eta' = (\eta \setminus \{ \Sigma[A(y_1, \ldots, y_{n(A)})] \} \cup \Sigma[\phi] \). We can assume w.l.o.g. that \( \{ z_1, \ldots, z_{n(A)} \} \cap \text{fV}(\phi) = \emptyset \) (if necessary, by an \( \alpha \)-renaming of existentially quantified variables), hence \( \gamma_0 \bowtie \phi \) and \( \gamma_0' = A(y_1, \ldots, y_{n(A)}) \), where \( \gamma_0' \bowtie (C_0, I_0, I_0, \psi, \Gamma) \bowtie \{ \psi \} \) and \( \gamma_0' \bowtie (C_1, I_1, I_1, \psi, \Gamma) \bowtie \{ \psi \} \), thus \( \gamma' \bowtie \phi \cdot A(y_1, \ldots, y_{n(A)}) \), where \( \gamma' = \gamma_0' \cdot \gamma' \). By the inductive hypothesis, the premiss \( \eta' \bowtie \{ \exists z_1 \ldots \exists z_{n(A)} \cdot \phi \cdot \phi(x_1, y_1, \ldots, y_{n(A)}/y_{n(A)}) \} \) \( L' \{ \psi \} \) is distinctive and, moreover, \( \gamma' \bowtie \exists z_1 \ldots \exists z_{n(A)} \cdot \phi \cdot \phi(x_1, y_1, \ldots, y_{n(A)}/y_{n(A)}) \), hence \( \{ \Sigma[\alpha] \}(\gamma') \cap \{ \Sigma[\alpha] \}(\gamma') = \emptyset \), for all \( \alpha \in \text{atoms}(\phi) \) and \( \alpha \in \text{atoms}(\phi) \). Since \( A(y_1, \ldots, y_{n(A)}) \) is precise on \( \{ \phi \cdot A(y_1, \ldots, y_{n(A)}) \} \), we have \( \{ \Sigma[A(y_1, \ldots, y_{n(A)})] \}(\gamma') = \bigcup_{\alpha \in \text{atoms}(\phi)} \{ \Sigma[\alpha] \}(\gamma') \). Since \( \{ \Sigma[A] \}(\gamma') = \{ \Sigma[A] \}(\gamma) \), for each \( \alpha \in \text{atoms}(\phi) \) \( \cup \).
\{A(y_1,\ldots,y_{\#(A)})\}\), we obtain that \(\langle\Sigma[\alpha]\rangle(\gamma) \cap \langle\Sigma[\alpha_2]\rangle(\gamma) = \emptyset\), for all \(\alpha_1, \alpha_2 \in \text{atoms}(\phi) \cup \{A(y_1,\ldots,y_{\#(A)})\}\), leading to the fact that \(\eta \vdash \{\phi \ast A(y_1,\ldots,y_{\#(A)})\}\) L \\(\{\psi\}\) is distinctive.

* (\(\lor\)) Let \(\eta \vdash \{\bigvee_{i=1}^{k} \phi_i \land \delta_i\}\) L \\(\{\bigvee_{i=1}^{k} \psi_i\}\) be the label of the root and let \(\gamma\) be a model of the preconditions of this triple. Then \(\gamma \models \phi \land \delta_i\), for some \(i \in [1, k]\). By the inductive hypothesis, the triple \(\eta \vdash \{\phi \land \delta_i\}\) L \\(\{\psi_i\}\) is distinctive, hence \(\langle\Sigma[\alpha]\rangle(\gamma) \cap \langle\Sigma[\alpha_1]\rangle(\gamma) = \emptyset\), for all \(\alpha_1, \alpha_2 \in \text{atoms}(\phi)\). Since \(\text{atoms}(\phi) = \text{atoms}(\bigvee_{i=1}^{k} \phi_i \land \delta_i)\), we obtain that \(\eta \vdash \{\bigvee_{i=1}^{k} \phi_i \land \delta_i\}\) L \\(\{\bigvee_{i=1}^{k} \psi_i\}\) is distinctive.

* (\(\land\)) and (\(\land\)): these cases are similar to (\(\lor\)).

* (\(C\)), (\(\ast\)), (\(U\)) and (\(C\)): these cases are trivial, because the preconditions and the environment does not change between the conclusion and the premises of these rules. \(\square\)

**Theorem 2** If \(\vdash \eta \vdash \{\phi\}\) L \\(\{\psi\}\) then \(\eta \vdash \{\phi\}\) L \\(\{\psi\}\).

**Proof.** For each axiom and inference rule in Fig. 5, with premises \(\eta_i \vdash \{\phi_i\}\) L \\(\{\psi_i\}\), for \(i = 1, \ldots, k, k \geq 0\), and conclusion \(\eta \vdash \{\phi\}\) L \\(\{\psi\}\), we prove that:

* (\(\ast\)) \(\vdash \eta \vdash \{\phi\}\) L \\(\{\psi\}\), if \(\vdash \eta_i \vdash \{\phi_i\}\) L \\(\{\psi_i\}\), for all \(i \in [1, k]\)

Let us show first that (\(\ast\)) is a sufficient condition. If \(\vdash \eta \vdash \{\phi\}\) L \\(\{\psi\}\) then there exists a cyclic proof whose root is labeled by \(\eta \vdash \{\phi\}\) L \\(\{\psi\}\) and we apply the principle of infinite descent to prove that \(\vdash \eta \vdash \{\phi\}\) L \\(\{\psi\}\). Suppose, for a contradiction, that this is not the case. Assuming that (\(\ast\)) holds, each invalid node, with label \(\eta_i \vdash \{\phi_i\}\) L \\(\{\psi_i\}\) and counterexample \(\gamma_i\), not on the frontier of the tree proof, has a successor, whose label is invalid, for all \(i \geq 0\). Let \(\text{preds}(\phi_i) = \{A_i(y_i),\ldots,A_i(y_i)\}\) be the set of predicate atoms from \(\phi_i\), for each \(i \geq 0\). Consequently, there exists a set of configurations \(\Gamma_i = \{\gamma_0,\ldots,\gamma_k\}\), such that \(\gamma_i = \gamma_0 \ast \cdots \ast \gamma_k\) and \(\gamma_i \vdash A_i(\gamma_j)\), for all \(j \in [1, k]\) and all \(i \geq 0\).

**Fact 1** For each \(i \geq 0\), either \(\Gamma_{i+1} \subseteq \Gamma_i\) or there exists \(j \in [1, k]\), such that \(\Gamma_{i+1} = (\Gamma_i \setminus \{\gamma_j\}) \cup \{\gamma_j \in [A(x_1,\ldots,x_{\#(A)})] \mid \gamma_j \subseteq \gamma_i, A(x_1,\ldots,x_{\#(A)}) \in \text{preds}(\phi_j(x_j/y_j))\}\), where \(A_i(\gamma_j) \leftarrow \exists x_j, \phi_i^{j}\) is a rule of the SIE and \(\phi_i^{j}\) is a symbolic configuration.

**Proof.** By inspection of the inference rules in Fig. 5b-e. The only interesting cases are:

* (\(\Rightarrow\)) in this case \(\Gamma_{i+1} \subseteq \Gamma_i\), because the models of the preconditions from the premises are subconfigurations of the model of the precondition in the conclusion,

* (\(LU\)) in this case \(\Gamma_{i+1}\) is obtained by replacing an element \(\gamma_j\) from \(\Gamma_i\) with a set of configurations \(\gamma_i\), such that \(\gamma_j \subseteq \gamma_i\) and \(\gamma_i\) is a model of a predicate atom from an unfolding of the predicate atom for which \(\gamma_j\) is a model. \(\square\)
For a configuration $\gamma' \in [[A_i(x'_j)]]$, we denote by $n(i, j)$ the minimum number of steps needed to evaluate the $\models$ relation in the given SID. Since $A(y_1, \ldots, y_{|\alpha|})$ is precise on $[\phi \ast A(y_1, \ldots, y_{|\alpha|})]$, we have $\langle\langle \Sigma|A(y_1, \ldots, y_{|\alpha|})\rangle\rangle(\gamma') = \bigcup_{\phi \in \text{atoms}(\phi)} \langle\langle \Sigma|\alpha\rangle\rangle(\gamma')$. Since $\langle\langle \Sigma|\alpha\rangle\rangle(\gamma') = \langle\langle \Sigma|\alpha\rangle\rangle(\gamma)$, for each $\alpha \in \text{atoms}(\phi) \cup \{A(y_1, \ldots, y_{|\alpha|})\}$, we obtain that $\langle\langle \Sigma|\alpha\rangle\rangle(\gamma') \cap \langle\langle \Sigma|\alpha_2\rangle\rangle(\gamma) = \emptyset$, for all $\alpha_1, \alpha_2 \in \text{atoms}(\phi) \cup \{A(y_1, \ldots, y_{|\alpha|})\}$, leading to the fact that $\eta \models [\phi \ast A(y_1, \ldots, y_{|\alpha|})]$, L $\{\psi\}$ is distinctive. Let $m_i$ be the multiset of numbers $n(i, j)$, for all $j \in [1, k_i]$ and $i \geq 0$. By Fact 1, the sequence of multisets $m_0, m_1, \ldots$ is such that either $m_0 = m_{i+1}$ or $m_i > m_{i+1}$, where the Dershowitz-Manna multiset ordering $\prec$ is defined as $m \prec m'$ if and only if there exist two multisets $X$ and $Y$, such that $X \neq \emptyset$, $X \subseteq m', m = (m' \setminus X) \cup Y$, and for all $y \in Y$ there exists some $x \in X$, such that $y < x$. By the fact that the cyclic proof tree is a cyclic proof, the infinite path goes infinitely often via a node whose label is the conclusion of the application of (LU). Then the infinite sequence of multisets $m_0, m_1, \ldots$ contains a strictly decreasing subsequence in the multiset order, which contradicts the fact that $\prec$ is well-founded.

Let $\sigma[w](\gamma) \overset{def}{=} \{\gamma' \mid \gamma \rightarrow_{w} \gamma'\}$, where the relation $\gamma \rightarrow_{w} \gamma'$ is defined in Def. 7. We are left with proving $(\ast)$ for each type of axiom and inference rule in Fig. 5:

- (e) For each configuration $\gamma$, we have $\langle\langle \varepsilon \rangle\rangle(\gamma) = \{\varepsilon\}$ and $\sigma[\varepsilon](C, I, \rho, v) = \{(C, I, \rho, v)\}$.

- (f) In any model $(C, I, \rho, v)$ of $\phi$, we have $\sigma[[v(x_1), p_1, v(x_2), p_2]](C, I, \rho, v) = \emptyset$, because of the side condition $\phi \uparrow \langle x_1, p_1, x_2, p_2 \rangle$ (Def. 11).

- (⊥) Because the precondition has no models.

- (Σ) By an application of Def. 7.

- (1−) Let $\gamma = (C, I, \rho, v) \in [[\phi \ast (x_1, p_1, x_2, p_2)]]$ be a configuration. By Lemma 3, we have that $\eta \models \langle\langle \phi \ast (x_1, p_1, x_2, p_2)\rangle\rangle L \{\psi \ast (x_1, p_1, x_2, p_2)\}$ is distinctive. By the side condition $\Sigma[[x_1, p_1, x_2, p_2]] \in \eta \setminus \text{supp}(L)$, it follows that $\langle\langle x_1, p_1, x_2, p_2 \rangle\rangle(\gamma)$ is disjoint from the interpretation $\langle\langle \Sigma|\alpha\rangle\rangle(\gamma)$ of any alphabet symbol $\Sigma|\alpha| \in \text{supp}(L)$, hence the interaction $[[v(x_1), p_1, v(x_2), p_2]](C, I, \rho, v)$ does not occur in $\langle\langle \Sigma|\alpha\rangle\rangle(\gamma)$. By the inductive hypothesis, we have $\models \eta \models \{\psi\} L \{\psi\}$, which leads to the required $\models \eta \models \{\phi \ast (x_1, p_1, x_2, p_2)\}$.

- (1+) Let $\gamma = (C, I, \rho, v) \in [[\phi]]$ be a configuration and $\omega \overset{def}{=} \Sigma[\alpha_1] \cdots \Sigma[\alpha_k]$ be a finite concatenation of alphabet symbols from $\text{supp}(L)$. If $\alpha_i = \langle x_1, p_1, x_2, p_2 \rangle$, for some $i \in [1, k]$, then we have $\langle\langle \Sigma|\alpha_i\rangle\rangle(\gamma) = \emptyset$, because of the side condition $\phi \not\models (x_1, p_1, x_2, p_2)$. Then $\sigma[w](\gamma) = \emptyset \subseteq \{\psi\}$, for each $w \in \langle\langle \omega\rangle\rangle(\gamma)$. Otherwise, if $\langle x_1, p_1, x_2, p_2 \rangle$ does not occur on $\omega$, then $\sigma[w](\gamma) \subseteq \{\psi\}$, for each $w \in \langle\langle \Sigma|\alpha_i\rangle\rangle(\gamma)$, by the inductive hypothesis.

- (⇒) Let $\gamma = (C, I, \rho, v) \in [[\phi_1 \ast \phi_2]]$ be a configuration. Then there exist configurations $\gamma_i = (C, I, \rho, v) \in [[\phi_i]]$, for $i = 1, 2$, such that $\gamma = \gamma_1 \bullet \gamma_2$. Let $\gamma'_i$ be configurations such that $\gamma_i \subseteq \gamma_{3-i}$, and $\gamma'_i \models \phi_i \ast F(\phi_{3-i})$, for $i = 1, 2$. Because $\phi_i \ast F(\phi_{3-i})$ is a separated conjunction of interaction atoms, each of which is precise on $\Gamma$, it follows that $\phi_i \ast F(\phi_{3-i})$ is precise on $\Gamma$, thus $\gamma_i$ are unique, for $i = 1, 2$. Let $\phi_{3-i} \overset{def}{=} \gamma_{3-i}$, for $i = 1, 2$. Moreover, since $\eta = \Sigma[\phi_i \ast F(\phi_{3-i})]$, the only interactions in $\langle\langle \eta_i\rangle\rangle(\gamma_i)$ are the ones in $\langle\langle \eta_i\rangle\rangle(\gamma_i')$, hence $\langle\langle \eta_i\rangle\rangle(\gamma_i) = \langle\langle \eta_i\rangle\rangle(\gamma_i')$, for
Let $w \in \{L_1 \triangleright \circ_{\eta_1, \eta_2} L_2\}(\gamma)$ be a word. Then $w \downarrow (\eta_1, \eta_2) \in \{L_i\}(\gamma)$, for $i = 1, 2$, because $(\eta_1, \gamma) = (\eta_1, \gamma')$, we obtain $w \downarrow (\eta_1, \gamma) = w \downarrow (\eta_1, \gamma') \in \{L_i\}(\gamma')$, for $i = 1, 2$. Since, moreover, $\gamma' = \phi_1 \ast F(\phi_1, \phi_{k-1})$, by the inductive hypothesis we obtain that $\sigma[w \downarrow (\eta_1, \gamma')] \subseteq [\psi_1 \ast F(\phi_1, \phi_{k-1})]$, for $i = 1, 2$. We partition $w = w_1 w'_1 w''_1 \ldots w_k w'_k w''_k$, for some $k \geq 1$, into three types of (possibly empty) blocks, such that, for all $j \in [1, k]$, we have:

1. $w_j \in (\{\eta_1\}(\gamma') \setminus \{\eta_2\}(\gamma'))^*$,
2. $w'_j, w''_j \in (\{\eta_1\}(\gamma') \cap \{\eta_2\}(\gamma'))^*$, and
3. $w_j \in (\{\eta_2\}(\gamma') \setminus \{\eta_1\}(\gamma'))^*$.

If $\sigma[w](\gamma) = \emptyset$, there is nothing to prove. Otherwise, let $\gamma = (C, I, \rho', v) \in \sigma[w](\gamma)$ and $\rho_1 \overset{\text{def}}{=} \rho, \rho_1', \rho_1'' \ldots, \rho_k, \rho_k', \rho_k''$ be a sequence of state maps such that, for all $j \in [1, k]$, we have $(C, I, \rho'_j, v) \in \sigma[w_j](C, I, \rho_j, v)$, $(C, I, \rho''_j, v) \in \sigma[w_j^j](C, I, \rho''_j, v)$, $(C, I, \rho''_j, v) \in \sigma[w_j^j](C, I, \rho''_j, v)$, and $(C, I, \rho_{j+1}, v) \in \sigma[w_{j+1}](C, I, \rho''_j, v)$, if $j < k$, in particular. Let $\rho_{j, j}$, $\rho'_{j, j}$, $\rho''_{j, j}$ and $\rho'''_{j, j}$ be the restrictions of $\rho_j$, $\rho'_j$, $\rho''_j$ and $\rho'''_j$ to $C_j$, for $i = 1, 2$, respectively. We prove the following:

1. $\rho_{2, j} = \rho'_{2, j}$, for all $j \in [1, k]$, and
2. $\rho''_{2, j} = \rho''_{2, j}$, for all $j \in [1, k]$.

We prove the first point, the argument for the second point being symmetric. It is sufficient to prove that the state of the components with indices in $C_2$, which are the only ones $\rho_{2, j}$ and $\rho'_{2, j}$ account for, is not changed by $w_j$, for all $j \in [1, k]$.

Since $w_j \in (\{\eta_1\}(\gamma') \setminus \{\eta_2\}(\gamma'))^*$, the only interactions on $w_j$ are the ones from $\gamma''_j = \gamma_j \cdot \gamma_j'$ that do not occur in $\gamma''_j = \gamma_j \cdot \gamma_j'$, where $\gamma_j \subseteq \gamma_1$ and $\gamma_j, \gamma_j' \subseteq \gamma_1$.

It follows that the interactions occurring on $w_j$ are the ones from $\gamma_j$ that do not occur in $\gamma''_j$. Since $\gamma''_j \models \alpha \otimes \text{inter}(\phi_1) \otimes \text{inter}(\phi_2)) \models \alpha \models F(\phi_2, \phi_1)$ 1], the interactions occurring on $w_j$ must occur in some model $\tilde{\gamma}$ of a tight subformula of $\phi_1$. Hence, the interactions from $\tilde{\gamma}$ can only change the state of a component from $\tilde{\gamma}$. Since $\tilde{\gamma} \subseteq \gamma_1$ and $\gamma_1 \cdot \gamma_1'$ is defined, there can be no component indexed by some element of $\alpha_2$, whose state is changed by an interaction from $\tilde{\gamma}$, thus $\rho_{2, j} = \rho'_{2, j}$ (1). Consequently, we obtain two sequences of words and finite state maps:

1. $w_1, w'_1, w''_1, \ldots, w_k, w'_k, w''_k$ and $\rho_{1, 1}, \rho_{1, 1}', \rho_{1, 1}'', \ldots, \rho_{1, k}, \rho_{1, k}', \rho_{1, k}''$, where:
   - $(C_1, I_1, \rho_{1, j}', v) \in \sigma[w_1](C_1, I_1, \rho_{1, j}, v)$,
   - $(C_1, I_1, \rho_{1, j}', v) \in \sigma[w'_1](C_1, I_1, \rho_{1, j}, v)$ and
   - $(C_1, I_1, \rho_{1, j+1}, v) \in \sigma[w''_1](C_1, I_1, \rho_{1, j}', v)$, for all $j \in [1, k - 1]$, and
2. $w'_1, w''_1, \ldots, w'_k, w''_k$ and $\rho'_{2, 1}', \rho'_{2, 1}'', \ldots, \rho'_{2, k}', \rho'_{2, k}''$, where:
   - $(C_2, I_2, \rho'_{2, j}', v) \in \sigma[w'_1](C_2, I_2, \rho'_{2, j}, v)$,
   - $(C_2, I_2, \rho'_{2, j}', v) \in \sigma[w''_1](C_2, I_2, \rho'_{2, j}, v)$ and
   - $(C_2, I_2, \rho''_{2, j}, v) \in \sigma[w'_1](C_2, I_2, \rho''_{2, j}, v)$ and
is precisely closed.

Proof. Let \( \varphi_i \overset{\text{def}}{=} \phi_i \ast k_i \text{tree}_i(x_{i,j}) \ast k_i \ast \text{tseg}(x_{i,j},y_{i,j}) \) be symbolic configurations, where \( \phi_i \) is a predicate-free symbolic configuration and \( \text{tree}_i(x) \) is either \( \text{tree}_{idle}(x) \), \( \text{tree}_{idle}(x) \) or \( \text{tree}(x) \), for all \( j \in [1,\ell_i] \) and \( i = 1,2 \). We prove that \( \varphi_i \) is precise on \( \llbracket \varphi_2 \rrbracket \). Let \( \gamma = (C_i,I_i,p_i,v) \in \llbracket \varphi_2 \rrbracket \) be a configuration and suppose that there exist configurations \( \gamma' = (C',I',p',v), \gamma'' = (C'',I'',p',v) \), such that \( \gamma' \subseteq \gamma \), \( \gamma'' \subseteq \gamma \), \( \gamma' \models \phi_1 \) and \( \gamma'' \models \phi_1 \). Then there exist configurations \( \gamma'_0 \overset{\text{def}}{=} (C'_0, I'_0, p'_0, v), \ldots, \gamma'_{k_i} \overset{\text{def}}{=} (C'_{k_i}, I'_{k_i}, p'_{k_i}, v) \) and \( \gamma''_0 \overset{\text{def}}{=} (C''_0, I''_0, p''_0, v), \ldots, \gamma''_{k_i} \overset{\text{def}}{=} (C''_{k_i}, I''_{k_i}, p''_{k_i}, v) \), such that:

- \( \gamma' = \ast_{j=0}^{k_i} \gamma'_j \) and \( \gamma'' = \ast_{j=0}^{k_i} \gamma''_j \),
- \( \gamma'_0 \models \phi_1 \) and \( \gamma''_0 \models \phi_1 \),
- \( \gamma'_j \models \text{tree}_i(x_{i,j}) \) and \( \gamma''_j \models \text{tree}_i(x_{i,j}) \), for all \( j \in [1,k_i] \), and
- \( \gamma'_j \models \text{tseg}(x_{i,j},y_{i,j}) \) and \( \gamma''_j \models \text{tseg}(x_{i,j},y_{i,j}) \), for all \( j \in [k_i+1,\ell_i] \).

Since \( \phi_1 \) is a predicate-free symbolic configuration, we have \( C'_0 = C''_0 \), \( I'_0 = I''_0 \) and \( p'_0 = p''_0 \), thus \( \gamma'_0 = \gamma''_0 \). Next, for each \( j \in [1,k_i] \), we have \( C'_j = C''_j \), because these sets of components correspond to the vertices of the same tree, whose root is \( v(x_{i,j}) \) and whose frontier contains only indices \( c \in C'_j \cap C''_j \), such that \( \rho(c) \in \{ \text{leaf}_{idle}, \text{leaf}_{busy} \} \). Finally, for each \( j \in [k_i+1,\ell_i] \), we have \( I'_j = I''_j \), because these sets of interactions correspond to the edges of the same tree, whose root is \( v(x_{i,j}) \) and whose frontier contains \( v(y_{i,j}) \) together with indices \( c \in C''_j \cap C''_j \), such that \( \rho(c) \in \{ \text{leaf}_{idle}, \text{leaf}_{busy} \} \).

We obtain, consequently, that \( \gamma'_j = \gamma''_j \), for all \( j \in [1,\ell_i] \), leading to \( \gamma' = \gamma'' \). \( \square \)

### D Havoc Invariance Proofs from Section 6

In order to shorten the following proofs, we introduce the rule (\( \dagger \)) that allows us to remove a disabled interaction atom \( \alpha \) from the pre- and postcondition, the environment and the language if certain conditions hold.
Lemma 4 Using the notation in §5, the following rule is sound:

\[
\frac{\eta \setminus \{\alpha\} \rightarrow \{\phi\} \perp \{\psi\}}{
\eta \rightarrow \{\phi + \alpha\} \cup \{\psi + \alpha\} \cup \{\phi \perp \alpha\} \rightarrow \{\psi \perp \alpha\}
}
\]

Proof. We assume that \(\phi\) and \(\psi\) are two symbolic configurations, \(\eta\) is an environment and \(\alpha = \langle x_1, p_1, x_2, p_2 \rangle\) an interaction atom. Furthermore, \(\Sigma \alpha \in \eta \setminus \text{supp}(L)\) and \(\phi \uparrow \alpha\). Then we can apply the rule (LU) first and the rules (C), (Σ) and (↑) on the subtrees and obtain:

\[
\begin{align*}
(1) \quad & \eta \rightarrow \{\phi + \alpha\} \cup \{\psi + \alpha\} \rightarrow \{\phi \perp \alpha\} \rightarrow \{\psi \perp \alpha\} \\
(2) \quad & \eta \rightarrow \{\phi + \alpha\} \cup \{\psi + \alpha\} \rightarrow \{\phi \perp \alpha\} \\
(3) \quad & \eta \rightarrow \{\phi + \alpha\} \cup \{\psi + \alpha\} \\
(4) \quad & \eta \rightarrow \{\phi + \alpha\} \cup \{\psi + \alpha\}
\end{align*}
\]

Hence the rule can be derived from the rules in Fig. 5.

D.1 Havoc Invariance of the Predicate Atom \(\text{tree}(x)\)

The invariance of the predicate \(\text{tree}(x)\) is proven via the rules in Fig. 5. The proof is divided into subtrees labeled by letters. Backlinks are indicated by numbers and in each cycle in the proof tree the rule (LU) is applied at least once.
D.2 Havoc Invariance of the Predicate Atom $\text{tree}_{\text{idle}}(x)$

The invariance of the predicate $\text{tree}_{\text{idle}}(x)$ is proven via the rules in Fig. 5 and the proof is structured similar to the previous invariance proof.
D.3 Havoc Invariance of the Predicate Atom tree_idle(x)

The proof of the invariance of the predicate tree_idle(x) is similar to the previous proofs.

\[ \text{backlink to (2)} \]

\[ (1) \]

\[ \text{(10)} \]

\[ \text{(11)} \]

\[ \text{(12)} \]
Lastly, we prove the invariance of the predicate \( \text{tseg}(x, u) \) via the rules in Fig. 5.
