CommCSL: Proving Information Flow Security for Concurrent Programs using Abstract Commutativity

MARCO EILERS, ETH Zurich, Switzerland
THIBAULT DARDINIER, ETH Zurich, Switzerland
PETER MÜLLER, ETH Zurich, Switzerland

Information flow security ensures that the secret data manipulated by a program does not influence its observable output. Proving information flow security is especially challenging for concurrent programs, where operations on secret data may influence the execution time of a thread and, thereby, the interleaving between threads. Such *internal timing channels* may affect the observable outcome of a program even if an attacker does not observe execution times. Existing verification techniques for information flow security in concurrent programs attempt to prove that secret data does not influence the relative timing of threads. However, these techniques are often restrictive (for instance because they disallow branching on secret data) and make strong assumptions about the execution platform (ignoring caching, processor instructions with data-dependent execution time, and other common features that affect execution time).

In this paper, we present a novel verification technique for secure information flow in concurrent programs that lifts these restrictions and does not make any assumptions about timing behavior. The key idea is to prove that all mutating operations performed on shared data commute, such that different thread interleavings do not influence its final value. Crucially, commutativity is required only for an *abstraction* of the shared data that contains the information that will be leaked to a public output. Abstract commutativity is satisfied by many more operations than standard commutativity, which makes our technique widely applicable.

We formalize our technique in CommCSL, a relational concurrent separation logic with support for commutativity-based reasoning, and prove its soundness in Isabelle/HOL. We have implemented Comm-CSL in HyperViper, an automated verifier based on the Viper verification infrastructure, and demonstrate its ability to verify challenging examples.

CCS Concepts: • **Security and privacy** → **Logic and verification**; **Information flow control**; • **Software and its engineering** → **Formal software verification**.

Additional Key Words and Phrases: Commutativity, information flow, separation logic, concurrency

1 INTRODUCTION

Reasoning about information flow is important to ensure the confidentiality and integrity of data. The main goal of information flow security is to ensure the absence of *value channels*, that is, unwanted information flows through the *result values* computed by a program. For sequential

Authors' addresses: Marco Eilers, Department of Computer Science, ETH Zurich, Zurich, Switzerland, marco.eilers@inf.ethz.ch; Thibault Dardinier, Department of Computer Science, ETH Zurich, Zurich, Switzerland, thibault.dardinier@inf.ethz.ch; Peter Müller, Department of Computer Science, ETH Zurich, Zurich, Switzerland, peter.mueller@inf.ethz.ch.

175

$$t1, t2 := 0$$

$$\texttt{while } (t1 < 100) \texttt{ do } \{ \quad \| \quad \texttt{while } (t2 < h) \texttt{ do } \{$$
$$\qquad t1 := t1 + 1 \qquad \qquad \qquad t2 := t2 + 1$$
$$\} \qquad \qquad \qquad \qquad \qquad \}$$
$$s := 3 \qquad \qquad \qquad \qquad \quad s := 4$$
$$print(s)$$

Fig. 1. Example program: Timing channels become value channels.

programs, preventing value channels is conceptually simple, and is enabled by various existing type systems, program logics, and static analyses.

Depending on the application scenario, it is sometimes desirable to also prevent side channels that let an attacker obtain information about secret values indirectly, by observing parameters such as execution time. Preventing timing channels requires ruling out *all possible* sources of secret-dependent timing differences, which may depend on complex (but standard) hardware features like caching [Stefan et al. 2013] and speculative execution [Kocher et al. 2019; Lipp et al. 2018].

For sequential programs, proving the absence of value channels does not involve reasoning about timing. Thus, for the many scenarios where attackers cannot observe execution time at all (e.g., because the time when data is made public is decoupled from its computation time, which is common for batch processing) or not with sufficient precision (e.g., because execution time is hidden by a laggy network), simple reasoning about values is sufficient even on standard hardware.

However, this is not the case for concurrent programs with shared memory, where timing differences may result in value channels, which can be observed even by attackers that cannot observe execution time. The program in Figure 1 illustrates this problem.

Since no information about the secret value h is leaked to the output s through direct assignments or explicit control flow, each thread individually does not leak secret information. However, the execution time of the loop in the right thread depends on the value of h. This timing may affect the order in which the threads perform their assignment to s and, thus, which value is printed in the end. With a deterministic round-robin scheduler, the printed value leaks whether or not h is greater than 100. With a non-deterministic scheduler with a known probability distribution, the entire value of h is leaked probabilistically over multiple executions.

Information leaks that turn secret-dependent timing differences between threads into observable differences in the program state have long been recognized and are called *internal timing channels* [Volpano and Smith 1998]. Some existing solutions to this problem essentially eliminate shared memory between threads [Giffhorn and Snelting 2015; Vechev et al. 2010], which rules out many useful concurrent programming patterns. Others employ the techniques used to prevent standard timing channels. They forbid programs from including operations whose execution time depends on secret inputs [Murray et al. 2018; Sabelfeld and Sands 2000; Schoepe et al. 2020; Smith 2007]; in our example, they would reject the loop in the right thread. However, these techniques typically assume idealized hardware, where secret-depending branching is the only source of timing channels. In principle, they can be extended to standard hardware (e.g., one commonly prevents timing leaks through cache effects by forbidding all secret-dependent memory accesses [Jancar et al. 2022]). However, attempting to rule out *all* sources of timing differences requires information about compilers [Barthe et al. 2018] and the hardware on which code is to be executed [Andrysco et al. 2018; Cleemput et al. 2012], including hardware details manufacturers usually do not make public. As a result, it is virtually impossible to rule out timing channels with absolute certainty,

and thus, existing techniques *cannot rule out value channels in concurrent programs on standard hardware with certainty* either, since they require eliminating timing channels.

This work. We present a novel technique to prove the absence of value channels in concurrent programs that does *not* require reasoning about timing and, thus, is sound even in the presence of complex modern hardware features. In the common case that attackers cannot observe execution time, our technique enables conceptually simple information flow security proofs. When the attacker *can* observe timing to a degree, our technique can be combined with orthogonal techniques to try to rule out timing channels. Crucially, our technique is *modular*, i.e., it does not require explicit reasoning about thread interleavings, which would make verification computationally infeasible.

Our technique is based on the observation that different, secret-dependent interleavings of shared data mutations lead to the same final result if the mutating operations on shared data performed by different threads *commute*. For instance, the program above is (rightfully) rejected by our technique because the two assignments to the shared variable s do not commute; however, a variation of this example where the concurrent assignments are modified so that the left thread atomically assigns $s + 3$ to s, and the right thread assigns $s + 4$, is allowed: This program always increases s by 7, since the additions performed by the two threads commute and their (secret-dependent) order therefore does not influence the final result. Consequently, the adapted program contains no value channels.

Standard commutativity is a strong requirement that is satisfied only by a minority of operations of typical data structures. A key insight of our work is to restrict the commutativity requirement to those parts of the shared data that are required to be non-secret (or *public*). In our example, if s were secret (and consequently not printed at the end of the program) then the program would be information flow secure even though the assignments to s do not commute. To this end, we allow programmers to specify abstract views of shared data structures that contain the information that must remain public; all information not included in the abstraction is considered secret. It is then sufficient to prove *abstract commutativity*, that is, the order of two operations does not affect the abstract view of a data structure, but it may affect other aspects. For instance, two list-append operations commute under a set- or length-abstraction of the list, but do not commute on the concrete list (unless they append the same element).

Many programs expose public views on data structures that contain secret information. One common case is maintaining *internal data* that intentionally contains secret information (e.g., an employee database with individual salaries), and then exposing only a part of it (e.g., average salaries for reporting purposes); in this case, our technique allows specifying the projection to averages as the abstraction. Another common case is maintaining data that is *intended* to be public, but inadvertently tainted with secret information as a result of timing differences in a concurrent computation. For instance, a program may maintain a list of anonymized accounts, but the order in the list depends on the time it took to process the (secret) purchases of each account. In this case, we can use the multiset view of the list as the abstraction, and prove, for instance, that a sorted version of the list can be exposed without creating a value channel.

We formalize our technique as a relational concurrent separation logic with special constructs for reasoning about abstract commutativity of concurrent operations. This logic, CommCSL, enables *modular* proofs of information flow security, does not require reasoning about the timing behavior of threads, and is amenable to automation via SMT solvers. Our logic supports an expressive assertion language that can express concepts like value-dependent secrecy [Murray et al. 2018]. In addition to proving existing applications secure, the ideas behind our technique can also be used as a guiding principle when building concurrent applications that deal with secret data.

Contributions and outline. We make the following contributions:

- We show how commutativity can be used to prove information flow security in shared-memory concurrent programs. Our verification technique allows threads to flexibly manipulate secret data and does not require reasoning (nor make assumptions) about the timing of executions, such that it is directly applicable to programs running on standard hardware.
- We introduce the generalized notion of abstract commutativity, which requires operations to commute only relative to a user-defined and application-specific abstraction of the shared state. Abstract commutativity applies to a much wider set of common scenarios.
- We incorporate these techniques into CommCSL, a novel concurrent separation logic that enables modular proofs of information flow security using abstract commutativity.
- We formalize our logic and prove its soundness in Isabelle/HOL [Nipkow et al. 2002].
- We implement CommCSL in an automated, SMT-based verification tool called HyperViper based on the Viper verification infrastructure [Müller et al. 2016].
- We evaluate our tool by proving several challenging examples information flow secure.

This paper is organized as follows: Sec. 2 provides an informal overview of our technique. We formally define CommCSL in Sec. 3, and summarize its soundness proof in Sec. 4. In Sec. 5, we describe the implementation of the logic in HyperViper and use it to verify several challenging programming patterns. We discuss related work in Sec. 6 and conclude in Sec. 7.

2 OVERVIEW

In this section, we explain the central concepts behind CommCSL informally. We occasionally omit details for simplicity; the full logic with all features and checks will be presented in Sec. 3.

2.1 Problem Statement

Our goal in this paper is to prove that concurrent programs that have both high-sensitivity (high) and low-sensitivity (low) inputs do not leak information about the high inputs in their low output values[1]. Formally, this can be expressed as (termination-insensitive) *non-interference* [Volpano et al. 1996], a 2-safety *hyperproperty* [Clarkson and Schneider 2010], i.e., a property of pairs of finite execution traces of the program ($s(x)$ denotes the value of variable x in program store s):

Definition 2.1. A program c with a set of input variables I and output variables O, of which some subsets $I_l \subseteq I$ and $O_l \subseteq O$ are low, satisfies non-interference iff for all s_1, s_2 and s'_1, s'_2, if $\forall x \in I_l. s_1(x) = s_2(x)$ and $\langle c, s_1 \rangle \rightarrow^* \langle \texttt{skip}, s'_1 \rangle$ and $\langle c, s_2 \rangle \rightarrow^* \langle \texttt{skip}, s'_2 \rangle$, then $\forall x \in O_l. s'_1(x) = s'_2(x)$.

This definition captures value channels. As explained in the introduction, timing side channels can potentially be prevented using orthogonal techniques if necessary.

The example in Fig. 2 illustrates this property. The targetSize procedure determines the size of the target audience for a marketing campaign as the number of people in a given array of households that satisfy certain criteria. The procedure uses two worker threads that each iterate over half of the households, determine the number of target persons in each household, and add that number to a shared counter c. We assume that the number of households in the input as well as how many members of a household are in the target audience (i.e., the result value of countTargets) is low. However, the execution time of countTargets may depend on high data for a variety of reasons. For example, the implementation in Fig. 2 consists of a simple look-up in a hash map. Nevertheless, its execution time may depend on hash collisions with other keys whose presence in the map is secret. Our goal is to prove that the output of targetSize is low.

---

[1]We limit our presentation to two security labels, high and low, instead of the general case of having arbitrary lattices of labels; however, techniques for verifying information flow security with two levels can be used to verify programs with arbitrary finite lattices by performing the verification multiple times, once for every element of the lattice.

```
procedure targetSize(households) {          procedure worker(households, f, t, c) {
  // ensures: c is low                        for (i in f..t-1) {
  n := |households|                             targets := countTargets(household[i])
  c := createCounter(0)                         atomic:
  (worker(households, 0, n/2, c)                  c.add(targets)
   ||                                          }
  worker(households, n/2, n, c))            }
  return c
}                                           procedure countTargets(household: map[str, int]) {
                                              return household["nAdults"]
                                            }
```

Fig. 2. Example program: Multiple threads add values to a shared counter. households is an array containing customer data per household. The result of countTargets is low, but its execution time may be high, here due to potential hash collisions with other keys whose presence in the households map is secret.

2.2 Commutativity-Based Information Flow Reasoning

The high data accessed in countTargets may affect the execution time of the procedure and, thereby, the thread schedule and the intermediate values of the shared counter c. Due to this internal timing channel, the value of c *must* be considered high *during* the execution of the worker threads. However, *after* the worker threads have terminated, the counter value can safely be considered to be low for two reasons. First, as per our assumption, each individual value added is low. Second, although high information may have affected the *order* in which different values are added to the counter, this order does not affect the final counter value because the updates *commute*. This example illustrates the central insight of this paper: *Internal timing channels do not affect the final outputs of a program if the modifications performed by different threads commute.*

Building on this insight, we present a verification technique for proving non-interference of concurrent programs. This technique enforces four central properties (in addition to standard checks that are sufficient to ensure non-interference for sequential programs):

(1) Low initial value: For any shared data structure that is to be modified by multiple threads, the initial contents of the data structure is low.
(2) Number of modifications is low: High data does not influence if or how often a thread performs a modification of the shared data at any given point in the program.
(3) Modification arguments are low: Any data inserted into the shared data structure is low.
(4) Commutativity: The modifications performed by different threads commute with each other.

These four properties are sufficient to ensure that, for any pair of executions of the verified program with identical low but potentially different high inputs, the *final* value of any shared data structure is identical after all threads have finished modifying it.

To understand why, consider two such executions: By property (1), the data structure will have the same contents in both executions when it is initially shared. By property (2), in both executions, the *same number of modifications* will be performed on the data structure (e.g., in our example, the same number of values will be added). By property (3), for each modification in the first execution, there is a matching modification in the second execution that uses the same arguments (e.g., if the value x has been added to the counter n times in the first execution, then x is also added n times in the second execution); the only difference is in the order of modifications. However, by property (4), reordering the modifications leaves the final value unchanged. Consequently, the final contents of the data structure are the result of performing the same commutative operations with the same arguments on the same initial value, and must therefore be identical.

The basic idea behind our verification technique is thus to prove these four properties for every shared data structure, which then allows us to treat the *final* contents of the data structure after

```
procedure targets(households) returns (res) {          procedure worker(households, f, t, m) {
  // ensures: res is low                                 for (i in f..t−1) {
  n := |households|                                        adr, rsn := select(household[i])
  m := createMap()                                         atomic:
  (worker(households, 0, n/2, m)                             m.put(adr, rsn)
    ||                                                     }
  worker(households, n/2, n, m))                         }
  res := sort(toList(keys(m)))
}
```

Fig. 3. Example program: Multiple threads add values to a shared map.

concurrent modifications have finished to be low (whereas all *intermediate* values read during concurrent modification will always have to be treated as high). In the rest of this section, we will build on and expand on this basic idea: We will relax some of the properties to make our verification technique more complete and more widely applicable, and we will explain how we check each of the four properties on the program to be verified.

2.3 Abstract Commutativity

The methodology described so far is sound, but not sufficiently complete: First, it requires that *all* information stored in shared data structures must be low after all concurrent modifications, even information that is never leaked to a public output. Second, most mutating operations on common data structures do not commute and therefore do not satisfy our property (4) above.

As an example, consider a variation of the previous example in Fig. 3: Now, the shared data structure is a map, and each worker extracts a key-value pair per household, where the address (the key) is low, but the reason why an address was selected (the value) is high. Here, different invocations of put do not always commute: If two threads put the same key but different values, then the later put-operation will "win", and its value will overwrite the previous value for said key. As a result, the final contents of the map allow an observer to conclude which put happened later and, thereby, draw conclusions about the secrets that cause the different thread interleavings.

However, the procedure targets does not actually return the entire contents of the map, but only (a sorted list representation of) its key set. Therefore, differences in values do not affect targets' public output. Scenarios where programs expose some public views on the data they maintain are common in practice; to accommodate them, we allow programmers to define an *abstract view* of the shared data structure that is guaranteed to be low and, therefore, allowed to be leaked. By focusing on the relevant part of a shared data structure, we no longer have to demand commutativity of all concurrent modifications, but only *abstract commutativity*: commutativity modulo the abstract view. That is, we require that, if the *abstract* view of the shared data is low initially, then switching the order of any two modifications on the data does not affect the final *abstract* view of the data structure. This is fulfilled for the example in Fig. 3: Different put-operations do not commute w.r.t. the entire contents of the map, but they do commute w.r.t. the map's key set. Note that our abstraction does *not* simply abstract away implementation details of the data structure to get a logical view of the data, but instead intentionally abstracts away all parts of the data structure that may contain high data, including otherwise vital parts of the data (like here the values in a map).

Since we require only the abstract view of shared data to be low, we no longer have to demand that *all* arguments of map modifications have to be low: arguments that do not affect the abstract view of the data may contain high information. In the example, this allows us to insert high *values* into the map as long as all inserted *keys* are low. We use preconditions on mutating operations to specify which arguments must be low to ensure that the abstract view of the data remains low.

To summarize, we relax our four central properties as follows:

(1) Low initial *abstract* value: The abstract view of the shared data is low when first shared.
(2) Number of modifications is low: Unchanged[2].
(3) Modification arguments fulfill sufficient precondition: (a) Arguments of each modification fulfill a given precondition that (b) ensures that the abstract view of the shared data remains low after the modification.
(4) *Abstract* Commutativity: All modifications commute w.r.t. the abstract view of shared data.

By enabling these relaxations, abstract commutativity allows us to handle many more practical examples than standard commutativity, at the cost of weakening the resulting guarantee: Now we may consider only the abstraction of the final value of the data structure to be low after concurrent modifications have been performed (all its intermediate values read during concurrent modification must still be treated as high). To demonstrate that the relaxed conditions are sufficient, we can make a similar argument as before to argue that the abstract view of each shared data structure will be low after all concurrent modifications on it have been performed.

2.4 Resource Specifications

Proving our four properties directly on the level of heap-manipulating programs is difficult: For example, showing that two modifications commute (even without an abstraction) on actual program states means proving the equivalence of two programs that perform the modifications in a different order. If these operations contain complex steps like pointer arithmetic, loops, or memory allocation (all of which might potentially be performed in real implementations of put methods on maps), it is non-trivial to even define equivalence of states (e.g., because of non-deterministic allocation), and even more difficult to perform such proofs [Koskinen and Bansal 2021].

We therefore do not check all of our properties *directly* on program states, but instead on pure mathematical values. We exploit the fact that, when working with separation logics, it is standard practice in the specification of data structures to use separation logic *predicates* that relate the contents of data structures to pure values [Parkinson and Bierman 2005]. For example, linked lists are typically specified using a predicate of the form $list(p, s)$, where p is the pointer to the start of the list and s is a mathematical sequence describing the contents of the list; $list(p, s)$ holds iff p points to a list whose contents are s. Methods that manipulate the data structures are then specified in terms of those pure values: for example, an append method of a list implementation would usually require the list predicate in its precondition for some abstract value s, and return it in its postcondition, with an abstract value that was extended by the appended value. We will assume that such specifications exist, and require additionally that the predicate uniquely determines the abstract value, which is typically the case for existing definitions. Thus, we exploit existing standard verification constructs to map our used data structures to pure values, enabling us to check properties (3) and (4) on the level of those pure values.

To check property (4), we must first determine which modifications are performed on the shared data. Scanning the entire program for such modifications would not be modular. Instead, we associate a set of legal operations with a data structure when we share it, and subsequently enforce that all modifications that are performed by any thread correspond to one of the legal operations.

We express these aspects of a shared data structure using a novel specification construct: a *resource specification* declares a *pure* data type (for the contents of a data structure in the standard separation logic style), an *abstract view* in the form of an abstraction function α that maps a value of the pure data type to another mathematical value that characterizes which aspects of the data

---

[2]One could in principle restrict property (2) to modifications that affect the abstract view of the shared data structure. We omit this optimization since operations that do not affect the abstract value commute trivially and, therefore, do not complicate verification significantly.

$$type_{\mathrm{MK}}(v) \equiv K \rightharpoonup V \qquad\qquad\qquad type_{\mathrm{MD}}(v) \equiv K \rightharpoonup V$$

$$\alpha_{\mathrm{MK}}(v) \equiv dom(v) \qquad\qquad\qquad \alpha_{\mathrm{MD}}(v) \equiv v$$

$$SharedActions_{\mathrm{MK}} \equiv \{\mathrm{P{\scriptstyle UT}}\} \qquad\qquad SharedActions_{\mathrm{MD}} \equiv \emptyset$$

$$UniqueActions_{\mathrm{MK}} \equiv \emptyset \qquad\qquad UniqueActions_{\mathrm{MD}} \equiv \{\mathrm{P{\scriptstyle UT}}_1, \mathrm{P{\scriptstyle UT}}_2\}$$

$$f_{\mathrm{P{\scriptstyle UT}}}(v, \langle key, val \rangle) \equiv m[k \mapsto v] \qquad\qquad f_{\mathrm{P{\scriptstyle UT}}_i}(v, \langle key, val \rangle) \equiv m[k \mapsto v]$$

$$pre_{\mathrm{P{\scriptstyle UT}}}(\langle key, val \rangle) \equiv \mathrm{Low}(key) \qquad\qquad pre_{\mathrm{P{\scriptstyle UT}}_i}(\langle key, val \rangle) \equiv \mathrm{Low}(key) \wedge \mathrm{Low}(val)$$

$$\wedge\; key \in range_i$$

Fig. 4. Left: Complete resource specification for the map example in Fig. 3, where K and V are the types of the keys and values in the map, respectively. Right: Alternative resource specification that allows two different threads to perform updates only in their own range of keys, which does not overlap with that of the other thread. We assume here that $range_1$ and $range_2$ do not overlap.

structure ultimately have to be public, and a set of *actions* that may be performed on the shared data structure. Each action comes with a function that defines how the action modifies the (pure) value of the data structure and with a (relational) precondition that restricts the arguments of the action such that, after performing the action, the abstract view of the data structure remains low.

Fig. 4 (left) shows a resource specification for our map example: Its type is a partial mapping from keys to values, its only legal action is Put, which updates the mapping, and the precondition of Put requires the argument key (but not the value) to be low using an assertion which we will formally introduce later. We will discuss later what it means for an action to be shared or unique.

Crucially, such a resource specification is independent of any specific client program and of any specific implementation of a map data structure; it can be used in any program that uses a shared map whose key set is low in the end, and it can be combined with any separation logic predicate for arbitrary map implementations (e.g., tree-based, list-based, hash-based, etc.).

A resource specification is *valid* if all its actions commute modulo its abstraction function (property (4)), and every action's precondition suffices to ensure the low-ness of the abstract view of the data (property (3b)); a formal definition of validity will follow in Sec. 3. The proof of validity has to be done only once per resource specification, and can be reused for different programs.

Crucially, our commutativity criterion (property (4)) enforces (abstract) *pairwise* commutativity of the actions in the resource specification, which avoids the (huge) effort of enumerating and comparing all possible interleavings of actions in a program: Pairwise commutativity of all actions is sufficient to permute the schedule of one execution into any possible schedule of a second execution with the same results (modulo abstraction).

2.5 Program Verification

Given a program and a resource specification that satisfies properties (3b) and (4), it remains to check properties (1), (2), and (3a) to prove that the program adheres to its resource specification.

Our logic for doing so, CommCSL, is a concurrent separation logic (CSL) [O'Hearn 2004]. To enable reasoning about low and high values, we phrase CommCSL as a relational logic that proves properties of two executions of the same program[3]. Like other concurrent separation logics, it requires showing that programs are data race free, by proving that each thread operates on its own partial heap separate from those of all other threads.

---

[3]While our technique could in principle also reason about more than two executions at a time, we focus only on pairs of executions, since that is sufficient for proving non-interference.

Invariants. As is standard in separation logics, heap locations that may be modified by different threads must be part of a shared resource, which is associated with an invariant that describes and constrains the resource's partial heap. In addition, we use the invariant to also map the shared data to some pure value, as described before; that is, the invariant is used to connect a heap-based shared data structure in the verified program to a value of the type defined in the resource specification. We denote the invariant as $I(v)$, where v denotes the pure value the invariant maps its heap to. As explained before, for a linked list, $I(v)$ would typically be some predicate $list(p, v)$ for some pointer p and sequence v. When initially sharing the resource, the invariant $I(v)$ must be established for some v. We check property (1) by enforcing that at this point, $\alpha(v)$ is low. The invariant itself must not express any low-ness constraint.

Atomic modifications. As explained in Sec. 2.4, we check property (4) on the level of abstract actions rather than concrete implementations. For this to be sound, *all* modifications of a shared data structure must be reflected by one of its actions, which we verify as follows.

In CommCSL (as in normal CSL), threads may modify the shared resource only in atomic-blocks. When entering such a block, they obtain the invariant, and on leaving the block, they have to re-establish the invariant and give it up again[4]. We use this mechanism to check that a program modifies a shared data structure only via the legal actions of the associated resource specification. Since a shared data structure may be modified only within an atomic-block, we can impose the following proof obligation: If, at the beginning of an atomic-block, $I(v)$ holds for some v, then at its end, $I(v')$ holds for some v' that is the result of applying one of the legal actions to the old value v, i.e., $v' = f_a(v, arg)$ for some action a and argument arg.

Guards. Let's turn to the remaining two properties, (2) and (3a). We could enforce property (2) by proving that there are no atomic-blocks under high guards (i.e., inside conditionals or loops whose conditions are high). However, this check would be overly conservative because it rules out implementations where the final value of a shared data structure is low even though intermediate states depend on a secret; our evaluation demonstrates that such examples occur in practice.

Therefore, we follow an alternative approach: we *record* all actions that are performed on the shared resource, and then check property (2) retroactively on the recorded actions when the resource is unshared (and all concurrent modifications must therefore have finished). We will follow the same approach for property (3a) and, thus, also record the arguments of each performed action, such that we can check action preconditions later.

For each action, we introduce a separation logic resource, which we call a *guard*, to record how often and with which arguments this action has been performed. Like other separation logic resources, guards can be transferred between methods, split into fractional parts such that they can be shared between threads, and subsequently be recombined. Guards are parameterized with a multiset of arguments, which records the arguments of the actions performed on a shared resource so far. These parameters reflect *which* actions have been performed, how often, and with which arguments, but not their order, which is not known due to the influence of scheduling.

When sharing a resource, we obtain a guard for each action that is legal according to the resource specification. The parameters of these guards are initially empty because no actions have been performed yet. Performing an action then imposes a proof obligation that some fraction of the respective guard is held, and adds the argument to the guard's multiset.

---

[4]*Obtaining* the invariant adds the partial heap described by the invariant to the current heap; *giving it up* amounts to checking that the current heap can be split into a partial heap that satisfies the invariant and a remainder, and removing the former. These logical steps are sometimes called *producing* and *consuming*, or *inhaling* and *exhaling* the invariant.

```
1   procedure targets(households) {
2      // ensures ∃v'. List(res, v') * Low(v')
3      n := |households|
4      m := createMap()
5      {Map(m, empty)}
6      ⇒ {Map(m, empty) * Low(dom(empty))}
7         // share
8         {guard_PUT(∅#, 1)}
9         ⇒ {guard_PUT(∅#, ½) * guard_PUT(∅#, ½)}
10        worker(households, 0, n/2, m) ||
11          worker(households, n/2, n, m)
12        {∃s₁, s₂. guard_PUT(s₁, ½) * PRE_PUT(s₁)*
13         guard_PUT(s₂, ½) * PRE_PUT(s₂)}
14        ⇒ {∃s. guard_PUT(s, 1) * PRE_PUT(s)}
15        // unshare
16     {∃v. Map(m, v) * Low(dom(v))}
17     return sort(toList(keys(m)))
18     {∃v, v'. Map(m, v) * Low(dom(v))*
19      List(res, v') * Low(v')}
20  }
```

```
1   procedure worker(households, f, t, m) {
2      // requires guard_PUT(∅#, ½)
3      // ensures ∃s'. guard_PUT(s', ½) * PRE_PUT(s')
4      {guard_PUT(∅#, ½)}
5      ⇒ {guard_PUT(∅#, ½) * PRE_PUT(∅#)}
6      for (i in f..t−1) {
7         {∃s'. guard_PUT(s', ½) * PRE_PUT(s')}
8         adr, rsn := select(household[i])
9         {∃s'. guard_PUT(s', ½) * PRE_PUT(s') * Low(adr)}
10        atomic:
11           {∃v. Map(m, v)}
12           m.put(adr, rsn)
13           {∃v. Map(m, v[adr ↦ rsn])}
14           ⇒ {Map(m, f_PUT(v, ⟨adr, rsn⟩))}
15        {∃s'. guard_PUT(s' ∪# {⟨adr, rsn⟩}#, ½)*
16         PRE_PUT(s') * pre_PUT(⟨adr, rsn⟩)}
17        ⇒ {∃s'. guard_PUT(s' ∪# {⟨adr, rsn⟩}#, ½)*
18         PRE_PUT(s' ∪# {⟨adr, rsn⟩}#)}
19     }
20  }
```

Fig. 5. Proof outline for the example from Fig. 3, verified against the resource specification from Fig. 4 (left). $I(v)$ is defined to be $Map(m, v)$, where Map is assumed to be a pre-existing separation logic predicate that relates the contents of a map m to a mathematical value v. Superscript # denotes multiset operations. $guard_{PUT}(s, f)$ denotes an f-fraction of the guard for action PUT with argument multiset s.

Guards allow us to check properties (2) and (3a) at the time when a resource is unshared. Unsharing requires all guards to be held, so that all performed actions are known. Unsharing then consumes the guards, so that no threads have the ability to perform any actions after this point. We show property (2) simply by proving that the *cardinality* of the argument multiset (i.e., how often each action has been performed) is low.

To ensure property (3a), we need to prove that each execution of an action satisfies its precondition. Since these preconditions can be relational (e.g., requiring an argument to be low), this proof must match the execution of an action in one run of the program with an execution in the other run. Proving preconditions retroactively when unsharing a resource gives the proof more freedom which executions of an action to match, which makes the proof technique more complete, and is possible because preconditions are expressed only over argument values and therefore state-independent.

Concretely, when unsharing, we require showing for each action a that there is a bijection between the elements of the multiset of arguments s in one program execution and its elements in the other execution (ensuring that their number is the same, i.e., the action has been performed the name number of times in both executions), such that each pair of elements fulfills the relational precondition of the action. We denote this fact via the assertion $PRE_a(s)$. For the PUT action of the map example, this means checking that there is a bijection that maps every key-value pair in the multiset in the first execution to a pair with the same key (but potentially a different value) in the second execution, since the precondition of PUT requires the keys to be low (but not the values).

2.6 Verification of the Map Example

We now show the entire proof for the map example from Fig. 3 in Fig. 5 (where we assume that our resource specification from Fig. 4 (left) is valid). The code indentation in targets indicates where the resource is shared. When sharing the resource, we must establish and give up the invariant $I(v)$, which we instantiate to $Map(m, v)$ (a separation logic predicate for a map, which relates the map's contents to the pure partial mapping v), and show that $\alpha(v)$ is low (property (1)), which is

the case because the map is empty in both executions. We then share the map, meaning that we obtain the guard for the PUT action (the only action in our resource specification), with an empty argument multiset. We split this guard into two parts and give each part to a worker.

Inside worker, we prove the loop invariant $\text{guard}_{\text{PUT}}(s', \frac{1}{2}) * PRE_{\text{PUT}}(s')$ for some s' (the argument multisets in proofs are typically existentially quantified, s.t. it is not necessary to track actual values in specifications). That is, we collect all arguments with which we performed the PUT action in s' (which also tracks how often we have performed the action) and show that each such argument satisfies PUT's precondition. As explained above, we are *required* to prove $PRE_{\text{PUT}}(s')$ only later when we unshare the resource, but in this example, it is convenient to maintain this property throughout. Initially, $PRE_{\text{PUT}}(s')$ holds trivially since s' is empty.

When the worker enters an atomic-block, it obtains the invariant $I(v)$ for some v (that is, $Map(m, v)$), and has to show at the end of the block that $I(v')$ holds, where $v' = f_{\text{PUT}}(v, arg)$ and $arg = \langle \text{adr}, \text{rsn} \rangle$. We assume that the specification of the put-method is sufficiently strong to prove this. Note that, since inside the atomic-block only the invariant (which cannot contain any low-ness constraints) is known about the shared data, if the code were to read shared data inside the block, the data would be implicitly treated as high. The proof rule for atomic-blocks adds arg to the argument multiset of the guard. Since we can prove at this point that $pre_{\text{PUT}}(arg)$ holds (i.e., adr is low), and $PRE_{\text{PUT}}(s')$ held for the old argument multiset s', we can show that the new argument multiset $s'' = s' \cup^\# \{arg\}^\#$ also fulfills $PRE_{\text{PUT}}(s'')$, and thus we maintain our loop invariant (which is also the postcondition of worker).

Back in targets, we unshare the resource, which consumes the guard. To do so, we recombine the fractional guards into a single one, whose argument multiset s is the union of the argument multisets of the individual guards, s_1 and s_2. Now we *must* prove that $PRE_{\text{PUT}}(s)$ holds, which follows from the postcondition of the workers, once the guards are combined. We may *now* assume that the abstract view $\alpha(v)$ of the *current* value v of the shared map is low, i.e., we know that the map's key set is low and have to prove the postcondition of targets, which states that the contents of the returned list are entirely low. We prove this using appropriate specifications of the procedures keys, toList, and sort that are called in the last line. For example, we use a postcondition for keys stating that if the key set of its input map is low, then the contents of the set it returns are low.

2.7 Unique Actions

The technique we presented so far requires that *all* modifications of the shared resource commute (modulo abstraction). However, in asymmetric thread collaborations, there are often actions that are performed by only one thread (e.g., in a producer-consumer scenario, a single producer thread might add values to a queue, whereas multiple consumer threads read and remove values from the queue); we call such actions *unique* actions (as opposed to *shared* actions that are performed by multiple threads). Since the order of applications of a unique action does not depend on scheduling, it is not necessary for unique actions to commute with themselves.

As an example, consider a variation of the map example where different workers work on different ranges of keys, and where both the keys and values PUT into the map are low. In such a program, it is never the case that a PUT of one thread overwrites a value that was PUT by a different thread, and therefore differences in timing do not lead to different final map contents.

We can support such examples by defining an alternative resource specification that declares n different PUT actions, where the precondition of each action PUT_i requires its argument key to be in a range separate from that of all other PUT actions (and both key and value to be low), as shown in Fig. 4 (right) for $n = 2$. We declare each of these actions to be unique, which means that they may be performed only by a single thread. For resource specification validity, unique actions are required to (abstractly) commute with all actions *except* themselves. That is, for example, PUT_1 has

$$c \quad ::= \quad x{:=}e \mid x{:=}[e] \mid [e]{:=}e \mid x{:=}\texttt{alloc}(e) \mid \texttt{skip} \mid c; c \mid \texttt{if } (b) \texttt{ then } \{c\} \texttt{ else } \{c\}$$
$$\mid \texttt{while } (b) \texttt{ do } \{c\} \mid c || c \mid \texttt{atomic } c$$

Fig. 6. Programming language. e ranges over integer-typed expressions, b over boolean-typed ones.

to commute with Put_2, but *not* with itself. We can now define α to be the identity function (i.e., use no abstraction at all), and prove programs secure even if they leak the entire map[5].

During program verification, we do not allow splitting the guards for unique actions, which ensures that only a single thread may perform them. In addition, guards for unique actions record the *sequence* of previous arguments, since their overall order is now no longer dependent on scheduling, and $PRE_a(s)$ for a unique action a and argument sequence s requires that the length of s is low, and for each index i, the values of $s[i]$ in both executions fulfill the action's precondition.

3 LOGIC

In this section, we formalize our technique in CommCSL, a relational concurrent separation logic with support for commutativity-based reasoning.

3.1 Language

We formalize our logic for an imperative, concurrent language with a mutable heap, whose commands are defined in Fig. 6. We assume as given some expression language for integer-typed expressions e and boolean expressions b, with the usual operations. Commands like assignments, conditionals, loops, parallel composition and sequential composition are standard; parallel compositions can be nested to create programs with more than two concurrent threads. Programs interact with the heap via read and store commands $x{:=}[e]$ and $[e]{:=}e$, as well as $x{:=}\texttt{alloc}(e)$, which allocates a single heap location, initializes its value to e, and assigns the resulting pointer to x; an extension to allocating multiple heap locations at once is straightforward. Like other CSLs, we formalize our logic for a single shared resource; threads can use the command $\texttt{atomic } c$ to atomically execute the command c while having access to the shared resource. However, our approach is not limited to having a single shared resource, and multiple resources are supported in our implementation analogous to other CSLs [Brookes 2004; O'Hearn 2004].

Program states have the form $\langle s, h \rangle$, where the store s is a map from names to (integer) values and the heap h is a partial map from locations (natural numbers) to integer values. Expression evaluation is deterministic and total, meaning in particular that references to uninitialized variables evaluate to some default value; we denote the value of expression e in store s by $[\![e]\!](s)$. Program configurations have the form $c, \langle s, h \rangle$ or abort; the latter represents a failed computation (which will never be reachable for verified programs). We use a small-step semantics with transitions of the form $c, \langle s, h \rangle \rightarrow c', \langle s', h' \rangle$ (for non-aborting steps). Our semantics is identical to that of other concurrent separation logics [Vafeiadis 2011]; its rules are shown in Eilers et al. [2022].

3.2 Resource Specifications

To simplify the formalization, we assume that our resource specification allows arbitrarily many unique actions but only *one* shared action, which is not restrictive, since one can merge multiple shared actions into one s.t. the argument selects which action to perform. Given this assumption, a resource specification for a resource value of type T has the form $\langle \alpha, f_{a_s}, F_{a_u} \rangle$, where α is the abstraction function of type $T \rightarrow T_\alpha$ (for some mathematical type T_α), f_{a_s} is the shared action, whose

---

[5]As an aside, this example also shows that the same data structure can have different resource specifications that have different demands and give different guarantees for the sensitivity of the final result.

precondition we denote as pre_{a_s}, and F_{a_u} is a family of unique actions, indexed by a finite type I; we will write f_{a_i} as a synonym for $F_{a_u}(i)$ for all $i \in I$, and denote the precondition of unique action f_{a_i} by pre_{a_i}. For each action a, f_a is a function of type $T \rightarrow T_{arg_a} \rightarrow T$, where T_{arg_a} is the argument type of the function. Both preconditions are *relational* preconditions on the argument of the action, i.e., they can require that (aspects of) the arguments are low. Thus, they are boolean-typed functions that take two arguments of type T_{arg_a}, where the two arguments denote the values of the argument in the two executions. Note that the preconditions *cannot* constrain the resource value itself; we explain this limitation and show how to work around it in Eilers et al. [2022] .

A resource specification is *valid* iff (A) every action's relational precondition is sufficient to preserve the low-ness of the abstract view of the resource value (property (3b) from Sec. 2), and (B) all *relevant* pairs of actions commute w.r.t. the abstract view (property (4) from Sec. 2). Relevant pairs of execution are the shared action paired with all actions including itself, and every unique action paired with all unique actions *except* itself; this reflects the fact that unique actions do not have to commute with themselves. Formally, we define validity as follows:

Definition 3.1. A resource specification $\langle \alpha, f_{a_s}, f_{a_u} \rangle$ is valid iff

(A) For all actions a, values v, v', and arguments arg, arg', if $\alpha(v) = \alpha(v')$ and $pre_a(arg, arg')$, then $\alpha(f_a(v, arg)) = \alpha(f_a(v', arg'))$.

(B) For all pairs of actions a, a' in $\{(a_s, a_s)\} \cup \{(a_s, a_i)|i \in I\} \cup \{(a_i, a_j)|i, j \in I \wedge i \neq j\}$ and all arguments arg, arg', if $\alpha(v) = \alpha(v')$ then $\alpha(f_{a'}(f_a(v, arg), arg')) = \alpha(f_a(f_{a'}(v', arg'), arg))$.

3.3 Extended Heaps

While our semantics works on ordinary program heaps h, we define our assertions and our logic on *extended heaps*, an enriched notion of heaps that represent both fractional permissions [Boyland 2003] and guards. An extended heap gh is a triple $\langle ph, gs, Gu \rangle$. ph is a standard *permission heap* [Bornat et al. 2005; Vafeiadis 2011] that can express *partial ownership* of a heap location, that is, a partial map from locations to pairs $\langle r, v \rangle$ of positive rational numbers of at most 1 and values v. For example, a permission heap that maps location l to $\langle v, \frac{1}{2} \rangle$ denotes a half permission to the heap location l, where value v is stored. Partial ownership allows multiple threads to concurrently *read* the same heap location, since reading a location requires only some positive permission amount, whereas modifying its value requires a permission of 1.

The shared action guard state gs and the family of unique action guard states Gu, where $Gu(i) = gu_i$, are specific to our technique. A guard for an action represents the *right* to perform that action, and it tracks the arguments with which said action has already been performed.

Every *unique* guard state gu_i is either $\bot$ or a *sequence* of argument values of the unique action a_i, and represents (when a resource is shared) the entire sequence of arguments with which the unique action has been performed. Since only one thread is allowed to execute a unique action, the order of executions is known and, thus, we can track the arguments in a sequence (as opposed to a multiset for shared actions). Moreover, unique guard states cannot be split or combined; that is, the thread that performs the unique action will have the entire sequence of argument values in its unique guard state, and all other threads will have a guard state of $\bot$. Thus, when adding two unique guard states, if one has a non-$\bot$ value, then the other must be $\bot$, otherwise addition is undefined (see Eilers et al. [2022] for the formal definition of guard heap addition).

Addition of guard heap families is defined pointwise. We write $\bot$ for guard heap families that are $\bot$ for all i, and $[i \mapsto gu]$ for a guard heap family whose value is gu for i, and $\bot$ for all other indices.

The *shared* guard state gs is either $\bot$ or a pair $\langle r, args_s \rangle$, where r is a positive rational number of at most 1 and $args_s$ is a *multiset* of argument values. The shared action may be performed by multiple threads if each of those threads has a positive fractional shared guard state. Thus, the

$$
\begin{aligned}
(s_1, gh_1), (s_2, gh_2) &\models emp &\Longleftrightarrow\quad & dom(ph_1) = \emptyset \wedge dom(ph_2) = \emptyset \\
(s_1, gh_1), (s_2, gh_2) &\models b &\Longleftrightarrow\quad & [\![b]\!]_{s_1} = 1 \wedge [\![b]\!]_{s_2} = 1 \\
(s_1, gh_1), (s_2, gh_2) &\models e_1 \mapsto^r e_2 &\Longleftrightarrow\quad & gh_1 = \{[\![e_1]\!]_{s_1} \mapsto \langle r, [\![e_2]\!]_{s_1}\rangle\} \wedge gh_2 = \{[\![e_1]\!]_{s_2} \mapsto \langle r, [\![e_2]\!]_{s_2}\rangle\} \\
(s_1, gh_1), (s_2, gh_2) &\models P * Q &\Longleftrightarrow\quad & \exists gh_1', gh_1'', gh_2', gh_2''.\, gh_1 = gh_1' \oplus gh_1'' \wedge gh_2 = gh_2' \oplus gh_2'' \wedge \\
& & & (s_1, gh_1'), (s_2, gh_2') \models P \wedge (s_1, gh_1''), (s_2, gh_2'') \models Q \\
(s_1, gh_1), (s_2, gh_2) &\models P \wedge Q &\Longleftrightarrow\quad & (s_1, gh_1), (s_2, gh_2) \models P \wedge (s_1, gh_1), (s_2, gh_2) \models Q \\
(s_1, gh_1), (s_2, gh_2) &\models \exists x.\, P &\Longleftrightarrow\quad & \exists v_1, v_2.\, (s_1[x \mapsto v_1], gh_1), (s_2[x \mapsto v_2], gh_2) \models P \\
(s_1, gh_1), (s_2, gh_2) &\models sguard(r, e) &\Longleftrightarrow\quad & gs_1 = \langle r, [\![e]\!]_{s_1}\rangle \wedge gs_2 = \langle r, [\![e]\!]_{s_2}\rangle \wedge \\
& & & Gu_1 = \perp \wedge Gu_2 = \perp \wedge dom(ph_1) = \emptyset \wedge dom(ph_2) = \emptyset \\
(s_1, gh_1), (s_2, gh_2) &\models uguard_i(e) &\Longleftrightarrow\quad & Gu_1 = [i \mapsto [\![e]\!]_{s_1}] \wedge Gu_2 = [i \mapsto [\![e]\!]_{s_2}] \wedge \\
& & & gs_1 = \perp \wedge gs_2 = \perp \wedge dom(ph_1) = \emptyset \wedge dom(ph_2) = \emptyset \\
(s_1, gh_1), (s_2, gh_2) &\models b \Rightarrow P &\Longleftrightarrow\quad & [\![b]\!]_{s_1} = [\![b]\!]_{s_2} \wedge [\![b]\!]_{s_1} \Rightarrow (s_1, gh_1), (s_2, gh_2) \models P \\
(s_1, gh_1), (s_2, gh_2) &\models \mathrm{Low}(e) &\Longleftrightarrow\quad & [\![e]\!]_{s_1} = [\![e]\!]_{s_2}
\end{aligned}
$$

Fig. 7. Assertion validity, where $gh_i = \langle ph_i, gs_i, Gu_i\rangle$.

shared guard state represents the (potentially partial) knowledge of the multiset of arguments with which the shared action has been executed so far. If r is 1, then $args_s$ denotes *all* arguments with which the action has been performed so far. On the other hand, two threads might have guard states with $r = \frac{1}{2}$ each and respective arguments $args_1$ and $args_2$, which each contain all arguments with which *that thread* has performed the shared action; then, in total, the action has been performed with the union of the multisets $args_1 \cup^\# args_2$. In general, when adding the arguments of two non-$\perp$ shared guard states, we take the union of the argument multisets, or that of one state if the other is $\perp$. The sum of two permission heaps $ph \oplus ph'$ is standard; intuitively, permission amounts are added (to a value of at most one) and values are unchanged (see Eilers et al. [2022]).

The sum of two extended heaps $\langle ph, gs, gu\rangle \oplus \langle ph', gs', gu'\rangle$ is defined as $\langle ph\oplus ph', gs\oplus gs', gu\oplus gu'\rangle$ iff the sums of all its components are defined. Finally, an extended heap can be *normalized* to a normal heap (which we denote by $norm(gh)$), by taking the permission heap and removing the permission amounts. That is, the normalized form of $\langle ph, gs, gu\rangle$ has the domain $dom(ph)$, and for each location $l \in dom(ph)$, the normalized heap has the value v s.t. $ph(l) = \langle \_, v\rangle$.

3.4 Assertions

Our assertion language is defined as follows:

$$P, Q ::= emp \mid b \mid e \mapsto^r e \mid P * Q \mid P \wedge Q \mid \exists x.\, P \mid sguard(r, e) \mid uguard_i(e) \mid b \Rightarrow P \mid \mathrm{Low}(e)$$

where P and Q range over assertions, r over positive rationals up to 1, b ranges over boolean and e over all expressions (including sequence- and multiset-typed expression, to describe the arguments of the two kinds of guard-assertions). Assertions are *relational*: their validity (Fig. 7) is defined over pairs of states, allowing them to express that expressions are low, i.e., equal in both states.

All basic separation logic assertions have their usual meaning, but applied to both states. For instance, standard points-to-assertions $e_1 \mapsto^r e_2$ represent a permission amount of r to location e_1 with a value of e_2. Separating conjunctions $P * Q$ hold if the extended heaps of *both* states can be separated into partial extended heaps that fulfill the conjuncts. Existentials are interpreted such that there can be *different* values for the quantified variable in both states, that is, $\exists x.\, e \mapsto^1 x$ expresses that e points to potentially different values in the two states (i.e., it might be high).

The assertion $sguard(r, e_a)$ represents a fractional amount r of the guard for the shared action with the multiset of arguments e_a; it holds in a state with store s iff the shared guard state is $\langle r, [\![e_a]\!]_s\rangle$, the permission heap is empty, and the unique guard state is $\perp$. Thus, a state fulfilling $sguard(r, e_a)$ can be added to any other state with a compatible shared guard state. Similarly,

$uguard_i(e_a)$ represents the guard for the unique action with index i and holds iff the sequence of arguments is $[\![e_a]\!]_s$, the shared guard state is $\bot$, and the permission heap is empty.

Finally, the relational assertion $\text{Low}(e)$ states that expression e is low. When used in an implication $b \Rightarrow \text{Low}(e)$, it can also express *value-dependent sensitivity* [Murray et al. 2018]. For example, a data structure might contain pairs of booleans and other values, where the boolean expresses the sensitivity of the other value; then, at runtime, clients could retrieve values from the data structure and e.g. output them to different channels based on their sensitivity.

We define $PRE(e)$ in terms of other assertion constructs[6]. $PRE_s(e)$, for the shared action, expresses that there is a bijection between the elements of the multiset-typed expression e in one state and the elements of e in the other state, s.t. each element in one state and its corresponding element in the other state together fulfill the relational precondition pre_{a_s} of the shared action. Formally, we define this notion recursively, by stating that there is an element x in e in one execution, and an element x' in e in the second execution, s.t. they together fulfill pre_{a_s}, and after removing both values from their respective multisets, $PRE(e)$ holds again for the resulting smaller multisets:

Definition 3.2. If the argument type of the shared action is T_{arg_s} and e is a multiset of T_{arg_s}, then

$$PRE_s(e) = \begin{cases} true & \text{if } e = \emptyset^{\#} \\ \exists x \in e.\ pre_{a_s}(x) \wedge PRE_s(e \setminus^{\#} \{x\}^{\#}) & \text{otherwise} \end{cases} \tag{1}$$

Since pre_{a_s} is relational, $PRE_s(e)$ is also relational. Moreover, since the existential quantifier may be interpreted differently in both executions, $PRE_s(e)$ expresses the existence of a bijection between the argument multiset, as discussed in Sec. 2.5.

For each unique action, we require that for the sequence of arguments e, the length of the sequence is low and the elements at each index fulfill the action's relational precondition:

$$PRE_i(e) = \text{Low}(|e|) \wedge \bigwedge_{j \in [0, |e|)} pre_{a_i}(e[j]) \tag{2}$$

We call an assertion *unary* if it does not restrict pairs of state relative to one another. That is, P is unary if for all s_1, gh_1, s_2, gh_2, if $(s_1, gh_1), (s_1, gh_1) \models P$ and $(s_2, gh_2), (s_2, gh_2) \models P$ then also $(s_1, gh_1), (s_2, gh_2) \models P$. Note that all assertions that do not syntactically contain any $\text{Low}(e)$-assertions (or $PRE(e)$-assertions that are defined in terms of $\text{Low}(e)$) are unary.

Finally, $noguard(P)$ states that the assertion P holds only for states whose guard states are all $\bot$.

3.5 Resource Contexts

A *resource context* combines a resource specification with an assertion that defines the shared heap data structure and maps it to a single mathematical value. That is, a resource context Γ has the form $\langle \alpha, f_{a_s}, F_{a_u}, I(x) \rangle$, where the *invariant* $I(x)$ is a parameterized unary assertion defining a valid state of the shared data structure with the value x. The value of x must be uniquely defined by the invariant, i.e., we require that if $(s_1, gh_1), (s_2, gh_2) \models I(v)$ and $(s_1, gh_1), (s_2, gh_2) \models I(v')$ then $v = v'$.

Given a resource context Γ, we can now formally connect the value of the guard states to the permission heap, using the notion of *consistency*. Consistency expresses that the value of the shared data structure on the heap is a possible result of applying the recorded actions with the recorded arguments in some order. We say that a state with shared guard state $\langle r, args_s \rangle$ and unique guard states $args_i$ for every unique action a_i is consistent with a context Γ from an initial value v_0 iff (1) its shared guard state is complete (i.e., $r = 1$), (2) $I(v)$ holds in the state, for some v, and (3) if FU_i is the sequence of applications of f_{a_i} with the arguments in $args_i$ (i.e., the first application in FU_i applies f_{a_i} with argument $args_i[0]$, etc.) for all i, and similarly, FS is the sequence of applications

---

[6]In our Isabelle formalization, $PRE(e)$ is actually defined as another assertion in the assertion language.

of f_{a_s} with the arguments in *some* permutation of $args_s$, then there is some interleaving of FS and all FU_i s.t. applying this interleaving to the initial value v_0 results in a final value of v.

3.6 Proof Rules

Our proof rules define a judgment of the form $\Gamma_\perp \vdash \{P\}c\{Q\}$, where $\Gamma_\perp$ is either $\perp$ or some resource context Γ. Intuitively, it expresses that a command c, when executed from two states that fulfill P and either no initial shared resource (if $\Gamma_\perp$ is $\perp$) or a shared resource fulfilling Γ, will not abort, and if they terminate, the resulting states will fulfill Q. The formal definition will follow in Sec. 4.

Figure 8 shows the most important proof rules of CommCSL, those for sharing and unsharing, as well as for atomic actions. All other rules, e.g., those for assignments, heap accesses, as well as sequential and parallel composition, and the rules for framing, existentials, and consequence, are standard and shown in Eilers et al. [2022] . The rules for loops and conditionals are similar to other relational logics but, crucially, do not require conditions to be low, and thus allow secret-dependent branching. If branch conditions are high, then postconditions have to be unary, which prevents indirect flows by making it impossible to have $\mathrm{Low}(e)$ in the postcondition.

The share rule captures both the share and unshare operation. It can be used only when the current resource context is $\perp$, i.e., when there is no shared data structure. Once the share rule is used, threads can access a shared data structure, and must use guards to justify their actions on it and record the arguments of the actions they perform. The share rule enforces that the new context's resource specification is valid, the resource invariant initially holds, and the value x of the resource is low modulo abstraction (property (1) from Sec. 2.3). The postcondition of the conclusion expresses that the invariant will hold again, with a new value x', whose abstract view is also low (and thus allows assuming that the abstract view of the *current* value of the shared data structure is low after concurrent modification has finished). In its premise, it gives access to the guards for the shared and the unique actions, initially with an empty argument multiset resp. argument sequences. It requires that after executing c, the guards are present again; in particular, the guard for the shared action must be present in its entirety, and therefore record all arguments of the shared action in its multiset. It also requires that PRE holds for both the sequences of arguments of the unique actions and the multiset of arguments of the shared action, meaning that we can now, retroactively, show that for each application of an action, the relational precondition was fulfilled (property (3a)).

There are two similar atomic-rules, one for the shared action and one for the unique actions. Both require that the resource context is not $\perp$, i.e., there is currently a shared resource, and that the guard for the respective action is initially present, which ensures that no thread can modify the shared data structure without holding the appropriate guard. For the shared action, any positive fraction of the guard is sufficient; for the unique actions, the whole unsplittable guard is required. In the postcondition of the conclusion, the guard records the new argument with which the action has been performed in its multiset or sequence, respectively. The rule's premise requires showing that, assuming the invariant holds initially for some value x_v (which, since the invariant is unary and thus no low-ness assumptions can be made, forces one to treat all shared data as high), it holds again after the atomic modification, with a new value that is the result of applying the respective action with the respective argument. We impose side conditions that require the assertions P and Q to not contain any guards (i.e., any guards held before the atomic block must be framed away using the frame rule), and that the invariant $I(x)$ is precise [O'Hearn et al. 2004], i.e., fulfilled by at most one sub-heap of any given heap. This condition is not limiting in practice, see Eilers et al. [2022] .

3.7 Limitations

CommCSL is not complete and can be extended in multiple ways. Its central limitations are that (1) it does not support information security proofs based on the absence of secret-dependent timing

$$\frac{\Gamma = \langle \alpha, f_{a_s}, f_{a_u}, I(x) \rangle \quad \Gamma \text{ is valid} \quad I(x) \text{ is unary and precise}}{\bot \vdash \{I(x) * \text{Low}(\alpha(x)) * P\} c \{\exists x'. I(x') * \text{Low}(\alpha(x')) * Q\}} \quad \text{(Share)}$$

where the premise above the middle line is:

$$\Gamma \vdash \{P * sguard(1, \emptyset^\#) * \textit{UniqueEmpty}\} c \{Q * sguard(1, x_s) * \textit{PRE}_s(x_s) * \textit{UniquePre}\}$$

$$\frac{\begin{array}{c} \Gamma = \langle \alpha, f_{a_s}, f_{a_u}, I(x) \rangle \quad I(x) \text{ is unary and precise} \\ x_v \notin fv(P) \cup fv(Q) \cup fv(I(x)) \quad x_s, x_a, x_v \notin mod(c) \quad noguard(P) \quad noguard(Q) \\ \bot \vdash \{P * I(x_v)\} c \{Q * I(f_{a_s}(x_v, x_a))\} \end{array}}{\Gamma \vdash \{P * sguard(r, x_s)\} \texttt{atomic } c \{Q * sguard(r, x_s \cup^\# \{x_a\}^\#)\}} \quad \text{(AtomicShr)}$$

$$\frac{\begin{array}{c} \Gamma = \langle \alpha, f_{a_s}, f_{a_u}, I(x) \rangle \quad I(x_v) \text{ is unary and precise} \\ x_v \notin fv(P) \cup fv(Q) \cup fv(I(x)) \quad x_s, x_a, x_v \notin mod(c) \quad noguard(P) \quad noguard(Q) \\ \bot \vdash \{P * I(x_v)\} c \{Q * I(f_{a_i}(x_v, x_a))\} \end{array}}{\Gamma \vdash \{P * uguard_i(x_s)\} \texttt{atomic } c \{Q * uguard_i(x_s \texttt{++}[x_a])\}} \quad \text{(AtomicUnq)}$$

Fig. 8. The central proof rules: the SHARE rule and the two ATOMIC rules implement all checks specific to our technique. Recall that the family of unique actions f_{a_u} is indexed by a (finite) type I; we denote the elements of I by $i_0, \ldots, i_n$. We write *UniqueEmpty* to abbreviate $uguard_{i_0}([]) * \cdots * uguard_{i_n}([])$, and *UniquePre* to abbreviate $\exists x_{i_0}, \ldots, x_{i_n}. uguard_{i_0}(x_{i_0}) * PRE_u(x_{i_0}) * \cdots * uguard_{i_n}(x_{i_n}) * PRE_u(x_{i_n})$. We use $fv(P)$ to denote the free variables in the assertion P, and $mod(c)$ for the set of variables modified by command c.

and thus cannot prove shared data low while it is being concurrently modified, (2) for programs that enforce an ordering between concurrent modifications, requiring commutativity of *all* actions is unnecessarily strong, (3) in specific situations, it would be sound to treat intermediate values of shared data as low, which CommCSL does not allow, and (4) it focuses on proving the output values of a program low and has no support for reasoning about the sensitivity of I/O (e.g., calls to print-procedures). Limitation (1) can be mitigated by combining CommCSL with existing logics like SecCSL and limitation (4) can be lifted via a simple extension of the logic (which we do in our implementation). Addressing limitations (2) and (3) is possible and left as future work.

In addition, CommCSL does not change the fact that concurrent code handling secret data has to be specifically designed to be secure. While with existing techniques, code must avoid all secret-dependent timing, our technique provides a new pattern for designing secure code based on commutativity. It is well-suited for programs that concurrently compute and return data in the presence of secrets, but other techniques are more complete for programs that output values during concurrent modifications, if their assumptions about timing are fulfilled.

4 SOUNDNESS

We have formalized CommCSL and proved it sound in Isabelle/HOL; here, we give a high-level overview of our proof (which is part of our artifact [Eilers et al. 2023] and available in the Archive of Formal Proofs [Dardinier 2023]). We build on Vafeiadis's soundness proof of CSL [Vafeiadis 2011], whose basic idea is to define a predicate $safe_n(P, c, Q)$, which (ignoring frames and resources for now) expresses that c, when executed for n steps from any state satisfying P, will not abort and, if it terminates after those n steps, will end up satisfying Q. Hoare triples are then defined to hold if $safe_n$ holds for all n. The same idea is used in the soundness proof for SecCSL [Ernst and Murray 2019], a concurrent separation logic for non-interference. SecCSL enforces that two executions with identical low inputs have the same control flow, s.t. the scheduler makes the same decisions in both executions, and therefore (in an idealized scenario) no internal timing channels exist. Since we allow high data to influence control flow and therefore timing and scheduling, we have to prove that, given two executions with identical low inputs with arbitrary, potentially *different* schedules (and different control flow per thread), the program's public output will be low.

As a result, the inductive argument used in existing proofs does not work for our setting. We solve this problem by separating the proof of safety from the proof of relational properties. We define a predicate $safe_n(\Gamma_\perp, c, \sigma, \Sigma)$, where $\Gamma_\perp$ is, as before, either $\perp$ or a resource context Γ, σ ranges over program states with extended heaps of the form $\langle s, gh\rangle$, and Σ is a set of such program states. Intuitively, $safe_n$ expresses (1) that executing c from σ, extended with a heap satisfying the resource invariant, for n steps will not abort, (2) that if it terminates, will result in *one* of the states in Σ (since, due to non-deterministic scheduling, there may be more than one final state), and (3) that if there is a shared resource, the state remains consistent w.r.t. some initial value v_0, i.e., that the state of the resource heap continues to be a possible result of applying the actions with the arguments specified in the guards to the initial value v_0. We provide the formal definition of safety in Eilers et al. [2022] . Crucially, this predicate makes a statement only about an individual execution.

We can now define the validity of a Hoare triple $\Gamma_\perp \models \{P\}c\{Q\}$ by stating that any two executions from a pair of states satisfying P are *safe* for an arbitrary number of steps, and *any* pair of final states from those two executions will fulfill Q. More precisely, the Hoare triple holds if there is some function Ψ that maps any initial state σ to a set of all possible final states Σ, and for any σ_1, σ_2 fulfilling P, all pairs of states in $\Psi(\sigma_1) \times \Psi(\sigma_2)$ fulfill Q:

Definition 4.1. $\Gamma_\perp \models \{P\}c\{Q\}$ holds iff there is some Ψ s.t. for all n, σ, if $\sigma, \sigma \models P$, we have $safe(n, \Gamma_\perp, c, \sigma, \Psi(\sigma))$, and for all $\sigma_1, \sigma_2, \sigma_1', \sigma_2'$, if $\sigma_1, \sigma_2 \models P$ and $\sigma_1' \in \Psi(\sigma_1)$ and $\sigma_2' \in \Psi(\sigma_2)$, then $\sigma_1', \sigma_2' \models Q$.

Our soundness proof uses a lemma stating, essentially, that the conditions (1)–(4) from Sec. 2 are sufficient to ensure the abstraction of the final shared value is low:

LEMMA 4.2. *If Γ is valid, $\alpha(v_0) = \alpha(v_0')$, v is consistent with v_0 through some sequence of shared action applications with argument multiset $args_s$ and unique action applications with argument sequences $args_i$ for all i, and similarly v' is consistent with Γ from v_0' through some sequence of shared action applications with argument multiset $args_s'$ and unique action applications with argument sequences $args_i'$ s.t. $PRE_s(args_s, args_s')$ and $PRE_i(args_i, args_i')$ for all i, then $\alpha(v) = \alpha(v')$.*

Here, condition (1) is $\alpha(v_0) = \alpha(v_0')$, conditions (2) and (3a) are expressed by the *PRE*-constraints, and the validity of Γ represents conditions (3b) and (4). Using this lemma, we prove:

THEOREM 4.3. (Soundness) *If $\Gamma_\perp \vdash \{P\}c\{Q\}$ then $\Gamma_\perp \models \{P\}c\{Q\}$.*

It follows that *intermediate* assertions proved at some program point will actually hold at this point for any pair of executions that reaches it, since it is always possible to cut off the rest of the program (and the proof) after such a point and apply the soundness theorem to the part of interest.

Thus, Hoare triples make true statements about pairs of program executions from an initial state with no shared resource (using $\widehat{gh}$ to range over extended heaps whose guard states are $\perp$ and whose permission heaps have full permission of every location in their domain):

COROLLARY 4.4. *If $\perp \vdash \{P\}c\{Q\}$ and $\langle s_1, \widehat{gh_1}\rangle, \langle s_2, \widehat{gh_2}\rangle \models P$, then c does not abort from any $\langle s_i, norm(\widehat{gh_i})\rangle$, and if $c, \langle s_i, norm(\widehat{gh_i})\rangle \to^* \mathrm{skip}, \langle s_i', h_i\rangle$ for some s_i', h_i and $h_i = norm(\widehat{gh_i}')$ for some $\widehat{gh_i}'$, then $\langle s_1', \widehat{gh_1}'\rangle, \langle s_2', \widehat{gh_2}'\rangle \models Q$.*

Finally, we can prove non-interference according to Def. 2.1 by proving that if all low input variables I_l have low values initially, then all low output variables O_l have low values in the end:

COROLLARY 4.5. *If $\perp \vdash \{P * \bigwedge_{x \in I_l} Low(x)\}c\{\bigwedge_{x \in O_l} Low(x)\}$ for some unary P and we have $\langle s_1, \widehat{gh_1}\rangle, \langle s_2, \widehat{gh_2}\rangle \models P * \bigwedge_{x \in I_l} Low(x)$, then if $c, \langle s_i, norm(\widehat{gh_i})\rangle \to^* \mathrm{skip}, \langle s_i', h_i\rangle$ for some s_i', h_i and $h_i = norm(\widehat{gh_i}')$ for some $\widehat{gh_i}'$, then $\langle s_1', \widehat{gh_1}'\rangle, \langle s_2', \widehat{gh_2}'\rangle \models \bigwedge_{x \in O_l} Low(x)$.*

Example	Data structure	Abstraction	LOC	Ann.	T
Count-Vaccinated	Counter, increment	None	44	46	10.15
Figure 2	Integer, add	None	129	95	10.90
Count-Sick-Days	Integer, add	None	52	45	13.67
Figure 1	Integer, arbitrary	Constant	29	20	1.52
Mean-Salary	List, append	Mean	80	84	14.10
Email-Metadata	List, append	Multiset	82	75	16.70
Patient-Statistic	List, append	Length	73	70	4.92
Debt-Sum	List, append	Sum	76	81	14.45
Sick-Employee-Names	Treeset, add	None	105	113	28.43
Website-Visitor-IPs	Listset, add	None	74	69	6.20
Figure 3	HashMap, put	Key set	129	96	10.37
Sales-By-Region	HashMap, disjoint put	None	129	104	12.37
Salary-Histogram	HashMap, increment value	None	135	109	13.78
Count-Purchases	HashMap, add value	None	137	109	11.73
Most-Valuable-Purchase	HashMap, conditional put	None	140	118	17.87
1-Producer-1-Consumer	Queue	Consumed sequence	82	88	3.23
Pipeline	Two queues	Consumed sequences	122	100	3.66
2-Producers-2-Consumers	Queue	Produced multiset	130	134	8.45

Table 1. Evaluated examples. We show the used data structure and the operation(s) we allow on it; for the last three examples, these are producing and consuming data. LOC are lines of code, not including specifications, Ann. are only specifications and proof annotations. T is the verification time in seconds, averaged over 5 runs.

5 IMPLEMENTATION AND EVALUATION

We have implemented our technique in HyperViper, an automated prototype verifier based on the Viper verification infrastructure [Müller et al. 2016], which is available as open source[7] and part of our artifact [Eilers et al. 2023]. HyperViper supports a richer language than the one used in this paper; in particular, instead of parallel composition commands, it allows dynamic thread creation using fork and join commands. Additionally, HyperViper supports *multiple* resources in a single program (which can be associated with different resource specifications).

HyperViper's level of automation is similar to existing automated verifiers: Users must provide method pre- and postcondition as well as loop invariants, in an assertion language that is similar to the one shown in this paper, and additionally have to declare resource specifications. Given these specifications and some annotations indicating key proof steps (e.g., which action is performed by an atomic block, and how to split and merge shared action guards; see Eilers et al. [2022] for an example), our tool automatically either verifies the program or indicates potential errors.

Internally, HyperViper encodes the validity constraints for all resource specifications as well as all other proof obligations imposed by our logic into the Viper intermediate language. To encode relational proof obligations, it uses a modular product program construction [Eilers et al. 2018], and combines it with existing encodings for concurrent programs [Leino and Müller 2009] in a sound way [Eilers et al. 2021]. Subsequently, it automatically verifies the generated program using one of Viper's backend verifiers and, ultimately, the Z3 SMT solver [de Moura and Bjørner 2008].

To demonstrate the practical usefulness of our verification technique, we have applied Hyper-Viper to a number of example programs that represent a variety of applications that concurrently manipulate shared data structures with secret data. The first 15 examples model specific applications, which spawn a number of worker threads to modify some shared data structure and subsequently

---

[7]https://github.com/viperproject/hyperviper

output the abstract view of the shared data structure. All examples explicitly model data structures containing different kinds of secret information (e.g., employee names and salaries, patient health data, or user activity on a website). Depending on the example, the data contains some low parts (e.g., whether or not a patient has been vaccinated), or some *aspect* of the secret data is low (e.g., the range of an employee's salary or the number of—otherwise secret—purchases they have made).

The last three examples model variations of general parallel programming patterns; here, we leave the data and computation abstract and show instead that the pattern is safe for *all* possible implementations where the individual roles do not directly leak secret data themselves. Table 1 shows, for each example, the total lines of code (not including whitespace), the number of lines used for annotations (i.e., specifications and proof annotations), and the verification time in seconds (measured on an 8-core AMD Ryzen 6850U with 32GB of RAM running Ubuntu 22.04 on a warmed-up JVM, averaged over five runs), as well as the used data structure and abstraction. The examples range from 29 to 140 lines of code, with a similar number of lines of specifications and proof annotations. Typical verification times are between 3 and 30 seconds. Thus, we conclude that commutativity-based reasoning can be automated efficiently. In the remainder of this section, we will discuss how different aspects of our technique allow us to verify the diverse set of examples.

Precise action definitions. For several examples, we can prove that the actions of all threads commute by carefully defining the legal actions. For example, while puts on maps with the same keys do not commute in general, in the example Salary-Histogram, each put *increments* the value for the key it is modifying (i.e., it increments the number of employees that fall within the specific salary range represented by the key). As a result, the updates commute. Similarly, the example Most-Valuable-Purchase iterates through purchase records and finds the highest price each user has ever paid: Here, each thread updates the map from user IDs to prices *only if* its purchase value is greater than the one currently in the map, which again results in commuting map updates.

Abstraction. In half of our examples, we use an abstraction to relax the commutativity requirement. For example, we compute a list of employee names and salaries but leak only the mean salary, and we collect a list of individual debts of a person but leak only the overall sum of their debt, not individual amounts or creditors. In example Email-Metadata, we abstract a list to a multiset to prove that we may leak the list after sorting it, which eliminates the secret-dependent item order. For the producer-consumer and pipeline examples, our abstraction does not return a view on the actual shared data, but on *ghost data* added for verification purposes: It does not state that the current data in the queue is low, which is not of interest, since the final state of the queue will always be empty, but that the sequence of *consumed* items is low (or, if there are multiple producers and consumers, the multiset view of that sequence). We also prove that the original version of Figure 1, where each thread sets the shared variable to a different value, is correct if said value is not leaked.

Resource specifications. The examples Sick-Employee-Names and Website-Visitor-IPs both add low values to sets (names of sick employees and IP addresses of website visitors, respectively), but use different set implementations. Since resource specifications abstract over concrete implementations of data structures, we can reuse the same resource spec for both examples.

Unique actions. In three examples, we use unique actions to exploit that some actions are performed only by one thread, and thus their order is unaffected by thread interleavings. In Sales-By-Region, different threads process data from different regions and add it to a shared map. Since the keys are region-specific, writes from different threads never conflict, and all updates commute. For the single-producer-consumer and pipeline examples, we exploit that there is only one producer and one consumer thread per shared queue to prove that the sequence of consumed values is low;

if there were multiple such threads (like in the multiple-producer-consumer example), the order of consumed items could be affected by secret data (and thus, only its multiset view is low).

Retroactive checking of action arguments. For the multiple producer-consumer example, the number of items consumed by each individual consumer depends on scheduling and therefore potentially high data. However, the *total* number of consumed items is low. Thus, checking that the number of performed consume actions is low *after* joining all threads, when the overall number of consume actions is known, allows us to prove this example secure. In the pipeline example, a thread produces low data into a first queue, a middle thread consumes this data, transforms it, and produces the result into a second queue, from where a third thread consumes it. While all threads are executing, the middle thread does not know that the data it reads from the first queue is low; we learn that only once the first queue is unshared. Thus, we would not be able to prove that the middle thread fulfills the second queue's precondition (that produced data is low) while all threads are running, but we *can* prove this precondition retroactively after unsharing the first queue.

High branches. Ca. half of our examples have secret-dependent timing due to branches on high data, and would thus be rejected by existing techniques, even if the attacker cannot observe timing.

6 RELATED WORK

Researchers have developed a plethora of type systems [Smith 2007], static analyses [Giffhorn and Snelting 2015], program transformations [Eilers et al. 2021], and program logics [Ernst and Murray 2019; Murray et al. 2018] to verify information flow security of concurrent programs, as well as multiple definitions of information flow security in this setting. Bisimulation-based properties [Focardi and Gorrieri 1995] and observational determinism [Zdancewic and Myers 2003] are properties of (sets of) traces, which assume that attackers can observe either low program variables or low events *during* the execution of the program, unlike our setting, where we assume that the attacker can observe *only* the public output of the program.

In our setting, the standard property is non-interference (which is what our logic guarantees), or weaker versions of it like *possibilistic* [Smith and Volpano 1998] and *probabilistic* [Sabelfeld and Sands 2000] non-interference, which guarantee, respectively, that secret data does not influence either the set of possible final outputs or their probabilities. Since possibilistic non-interference is too weak in practice, most existing techniques target either traditional or probabilistic non-interference. As discussed previously, they achieve this goal by entirely preventing secret-dependent execution time differences, under the idealized assumption that only high branches can lead to such differences [Eilers et al. 2021; Ernst and Murray 2019; Murray et al. 2018; Schoepe et al. 2020; Smith 2006, 2007]. These techniques can be applied to realistic settings by also preventing other sources of timing differences, but preventing them entirely on standard hardware is complex and requires strong assumptions about compilers and hardware. Our technique has the unique advantage of not requiring any reasoning about execution time in order to prove information flow security for output values, which makes it applicable independently of the used hardware. In settings where it is possible to precisely reason about execution time, COMMCSL *complements* existing techniques like SecCSL, and their relative completeness will vary from example to example (see Sec. 3.7). In principle, it is also possible to verify concurrent programs by rewriting them as sequential programs with non-deterministic scheduling and reasoning about them using techniques for sequential program. However, this approach would require explicitly considering the possible interleavings of all threads, which is not modular and does not scale in practice.

Most of the aforementioned techniques either do not support a mutable heap (e.g. [Murray et al. 2018; Smith 2007] or work for programs using locks to protect shared memory like our logic (e.g. [Eilers et al. 2021; Ernst and Murray 2019]), but some recent logics target more complex settings,

like fine-grained concurrency [Frumin et al. 2021] or relaxed memory models [Yan and Murray 2021]. We believe that our approach also extends to fine-grained concurrency, since the general idea of making sure that concurrent changes commute also applies in this setting.

Some techniques prevent unwanted information leakage (including leakage as a result of internal timing channels) at *runtime* by modifying the language runtime [Vassena et al. 2019], using information flow aware concurrency primitives [Stefan et al. 2012], or by transforming the original program [Russo et al. 2006]. Such techniques can allow running programs that would be rejected by static techniques, but at the cost of requiring specific runtimes or affecting execution performance.

There are many use cases of commutativity in the literature, particularly in the context of parallelization. CoMMCSL is the first technique that applies commutativity to information flow security in a concurrent setting. Benton et al. [2007] present a relational semantics that can *prove* that two operations commute if the locations one reads and the other modifies do not overlap, whereas we *use* commutativity of actions to prove noninterference for programs that perform these actions. Additionally, we consider cases where different threads *do* modify the same shared data, and thus Benton et al. [2007]'s reasoning principles do not apply to our setting.

Some related work uses commutativity coupled with an abstraction, in particular, it uses data structures that expose a kind of abstract view and define commutativity on the basis of such a view. For example, Kim and Rinard [2011] verify commutativity conditions in this context for specific data structures, and Bansal et al. [2018] automatically infer them. A lot of existing work *uses* commutativity with such abstractions to reason about or transform programs. For example, Golan-Gueta et al. [2015] use commutativity to parallelize sequential programs, Dimitrov et al. [2014] use commutativity to detect races in traces of concurrent programs, and Brutschy et al. [2017] use it to prove serializability. Kragl and Qadeer [2021] use commutativity to simplify reasoning about concurrent programs, and Pincus [2022] uses it to parallelize code automatically and additionally infers commutativity conditions. All of them use a notion of abstraction that is more akin to our pure (mathematical) representations of the data structure (e.g., a partial function for a hash map). Our notion of "abstract view" is an additional layer of abstraction on top of the pure mathematical value which is application-specific and intentionally abstracts away high data (e.g., by further abstracting the partial function to its domain). That is, existing work still uses and proves actual commutativity of data structure operations that ignores only internal implementation details of the data structure, whereas our abstractions allow operations that actually do not commute, if the result of this non-commutativity is not visible in public outputs. Thus, given a data structure and its pure representation, we may use different abstractions for different usages in the same program.

Finally, Farzan et al. [2023] abstract *statements* (e.g., a concrete assignment to a non-deterministic one) and reason about the commutativity of two abstracted statements *on concrete states*, whereas we abstract states and check whether actions commute w.r.t. these abstracted states.

7 CONCLUSION AND FUTURE WORK

We have presented CoMMCSL, a concurrent separation logic for verifying non-interference using abstract commutativity. We proved CoMMCSL sound and showed that it can be automated efficiently and is able to verify common examples. For future work, we plan to apply reasoning based on (abstract) commutativity to other settings like fine-grained concurrency, and to explore when it is sound to assume that shared data is low *while* threads are still performing concurrent updates.

ACKNOWLEDGMENTS

8 DATA AVAILABILITY STATEMENT

The formalization, implementation and evaluation of this paper are available in its artifact [Eilers et al. 2023]. The formalization is also available in the Archive of Formal Proofs [Dardinier 2023].

REFERENCES

Marc Andrysco, Andres Nötzli, Fraser Brown, Ranjit Jhala, and Deian Stefan. 2018. Towards Verified, Constant-time Floating Point Operations. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, CCS 2018, Toronto, ON, Canada, October 15-19, 2018*, David Lie, Mohammad Mannan, Michael Backes, and XiaoFeng Wang (Eds.). ACM, 1369–1382. https://doi.org/10.1145/3243734.3243766

Kshitij Bansal, Eric Koskinen, and Omer Tripp. 2018. Automatic Generation of Precise and Useful Commutativity Conditions. In *Tools and Algorithms for the Construction and Analysis of Systems - 24th International Conference, TACAS 2018, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2018, Thessaloniki, Greece, April 14-20, 2018, Proceedings, Part I (Lecture Notes in Computer Science, Vol. 10805)*, Dirk Beyer and Marieke Huisman (Eds.). Springer, 115–132. https://doi.org/10.1007/978-3-319-89960-2_7

Gilles Barthe, Benjamin Grégoire, and Vincent Laporte. 2018. Secure Compilation of Side-Channel Countermeasures: The Case of Cryptographic "Constant-Time". In *31st IEEE Computer Security Foundations Symposium, CSF 2018, Oxford, United Kingdom, July 9-12, 2018*. IEEE Computer Society, 328–343. https://doi.org/10.1109/CSF.2018.00031

Nick Benton, Andrew Kennedy, Lennart Beringer, and Martin Hofmann. 2007. Relational semantics for effect-based program transformations with dynamic allocation. In *Proceedings of the 9th International ACM SIGPLAN Conference on Principles and Practice of Declarative Programming, July 14-16, 2007, Wroclaw, Poland*, Michael Leuschel and Andreas Podelski (Eds.). ACM, 87–96. https://doi.org/10.1145/1273920.1273932

Richard Bornat, Cristiano Calcagno, Peter W. O'Hearn, and Matthew J. Parkinson. 2005. Permission accounting in separation logic. In *Proceedings of the 32nd ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL 2005, Long Beach, California, USA, January 12-14, 2005*, Jens Palsberg and Martín Abadi (Eds.). ACM, 259–270. https://doi.org/10.1145/1040305.1040327

John Boyland. 2003. Checking Interference with Fractional Permissions. In *Static Analysis, 10th International Symposium, SAS 2003, San Diego, CA, USA, June 11-13, 2003, Proceedings (Lecture Notes in Computer Science, Vol. 2694)*, Radhia Cousot (Ed.). Springer, 55–72. https://doi.org/10.1007/3-540-44898-5_4

Stephen D. Brookes. 2004. A Semantics for Concurrent Separation Logic. In *CONCUR 2004 - Concurrency Theory, 15th International Conference, London, UK, August 31 - September 3, 2004, Proceedings (Lecture Notes in Computer Science, Vol. 3170)*, Philippa Gardner and Nobuko Yoshida (Eds.). Springer, 16–34. https://doi.org/10.1007/978-3-540-28644-8_2

Lucas Brutschy, Dimitar K. Dimitrov, Peter Müller, and Martin T. Vechev. 2017. Serializability for eventual consistency: criterion, analysis, and applications. In *Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages, POPL 2017, Paris, France, January 18-20, 2017*, Giuseppe Castagna and Andrew D. Gordon (Eds.). ACM, 458–472. https://doi.org/10.1145/3009837.3009895

Michael R. Clarkson and Fred B. Schneider. 2010. Hyperproperties. *Journal of Computer Security* 18, 6 (2010), 1157–1210.

Jeroen V. Cleemput, Bart Coppens, and Bjorn De Sutter. 2012. Compiler mitigations for time attacks on modern x86 processors. *TACO* 8, 4 (2012), 23:1–23:20. https://doi.org/10.1145/2086696.2086702

Thibault Dardinier. 2023. Formalization of CommCSL: A Relational Concurrent Separation Logic for Proving Information Flow Security in Concurrent Programs. *Archive of Formal Proofs* (March 2023). https://isa-afp.org/entries/CommCSL.html, Formal proof development.

Leonardo Mendonça de Moura and Nikolaj Bjørner. 2008. Z3: An Efficient SMT Solver. In *Tools and Algorithms for the Construction and Analysis of Systems, 14th International Conference, TACAS 2008, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2008, Budapest, Hungary, March 29-April 6, 2008. Proceedings (Lecture Notes in Computer Science, Vol. 4963)*, C. R. Ramakrishnan and Jakob Rehof (Eds.). Springer, 337–340. https://doi.org/10.1007/978-3-540-78800-3_24

Dimitar Dimitrov, Veselin Raychev, Martin T. Vechev, and Eric Koskinen. 2014. Commutativity race detection. In *ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI '14, Edinburgh, United Kingdom - June 09 - 11, 2014*, Michael F. P. O'Boyle and Keshav Pingali (Eds.). ACM, 305–315. https://doi.org/10.1145/2594291.2594322

Marco Eilers, Thibault Dardinier, and Peter Müller. 2022. CommCSL: Proving Information Flow Security for Concurrent Programs using Abstract Commutativity. *CoRR* abs/2211.08459 (2022). https://doi.org/10.48550/arXiv.2211.08459 arXiv:2211.08459

Marco Eilers, Thibault Dardinier, and Peter Müller. 2023. *CommCSL: Proving Information Flow Security for Concurrent Programs using Abstract Commutativity (Artifact)*. https://doi.org/10.5281/zenodo.7709911

Marco Eilers, Severin Meier, and Peter Müller. 2021. Product Programs in the Wild: Retrofitting Program Verifiers to Check Information Flow Security. In *Computer Aided Verification - 33rd International Conference, CAV 2021, Virtual Event, July*

20-23, 2021, Proceedings, Part I (Lecture Notes in Computer Science, Vol. 12759), Alexandra Silva and K. Rustan M. Leino (Eds.). Springer, 718–741. https://doi.org/10.1007/978-3-030-81685-8_34

Marco Eilers, Peter Müller, and Samuel Hitz. 2018. Modular Product Programs. In *Programming Languages and Systems - 27th European Symposium on Programming, ESOP 2018, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2018, Thessaloniki, Greece, April 14-20, 2018, Proceedings (Lecture Notes in Computer Science, Vol. 10801)*, Amal Ahmed (Ed.). Springer, 502–529. https://doi.org/10.1007/978-3-319-89884-1_18

Gidon Ernst and Toby Murray. 2019. SecCSL: Security Concurrent Separation Logic. In *Computer Aided Verification - 31st International Conference, CAV 2019, New York City, NY, USA, July 15-18, 2019, Proceedings, Part II (Lecture Notes in Computer Science, Vol. 11562)*, Isil Dillig and Serdar Tasiran (Eds.). Springer, 208–230. https://doi.org/10.1007/978-3-030-25543-5_13

Azadeh Farzan, Dominik Klumpp, and Andreas Podelski. 2023. Stratified Commutativity in Verification Algorithms for Concurrent Programs. *Proc. ACM Program. Lang.* 7, POPL (2023), 1426–1453. https://doi.org/10.1145/3571242

Riccardo Focardi and Roberto Gorrieri. 1995. A Taxonomy of Security Properties for Process Algebras. *J. Comput. Secur.* 3, 1 (1995), 5–34. https://doi.org/10.3233/JCS-1994/1995-3103

Dan Frumin, Robbert Krebbers, and Lars Birkedal. 2021. Compositional Non-Interference for Fine-Grained Concurrent Programs. In *42nd IEEE Symposium on Security and Privacy, SP 2021, San Francisco, CA, USA, 24-27 May 2021.* IEEE, 1416–1433. https://doi.org/10.1109/SP40001.2021.00003

Dennis Giffhorn and Gregor Snelting. 2015. A new algorithm for low-deterministic security. *Int. J. Inf. Sec.* 14, 3 (2015), 263–287. https://doi.org/10.1007/s10207-014-0257-6

Guy Golan-Gueta, G. Ramalingam, Mooly Sagiv, and Eran Yahav. 2015. Automatic scalable atomicity via semantic locking. In *Proceedings of the 20th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, PPoPP 2015, San Francisco, CA, USA, February 7-11, 2015*, Albert Cohen and David Grove (Eds.). ACM, 31–41. https://doi.org/10.1145/2688500.2688511

Jan Jancar, Marcel Fourné, Daniel De Almeida Braga, Mohamed Sabt, Peter Schwabe, Gilles Barthe, Pierre-Alain Fouque, and Yasemin Acar. 2022. "They're not that hard to mitigate": What Cryptographic Library Developers Think About Timing Attacks. In *2022 IEEE Symposium on Security and Privacy, S&P 2022.*

Deokhwan Kim and Martin C. Rinard. 2011. Verification of semantic commutativity conditions and inverse operations on linked data structures. In *Proceedings of the 32nd ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI 2011, San Jose, CA, USA, June 4-8, 2011*, Mary W. Hall and David A. Padua (Eds.). ACM, 528–541. https://doi.org/10.1145/1993498.1993561

Paul Kocher, Jann Horn, Anders Fogh, Daniel Genkin, Daniel Gruss, Werner Haas, Mike Hamburg, Moritz Lipp, Stefan Mangard, Thomas Prescher, Michael Schwarz, and Yuval Yarom. 2019. Spectre Attacks: Exploiting Speculative Execution. In *2019 IEEE Symposium on Security and Privacy, SP 2019, San Francisco, CA, USA, May 19-23, 2019.* IEEE, 1–19. https://doi.org/10.1109/SP.2019.00002

Eric Koskinen and Kshitij Bansal. 2021. Decomposing Data Structure Commutativity Proofs with *mn*-Differencing. In *Verification, Model Checking, and Abstract Interpretation - 22nd International Conference, VMCAI 2021, Copenhagen, Denmark, January 17-19, 2021, Proceedings (Lecture Notes in Computer Science, Vol. 12597)*, Fritz Henglein, Sharon Shoham, and Yakir Vizel (Eds.). Springer, 81–103. https://doi.org/10.1007/978-3-030-67067-2_5

Bernhard Kragl and Shaz Qadeer. 2021. The Civl Verifier. In *Formal Methods in Computer Aided Design, FMCAD 2021, New Haven, CT, USA, October 19-22, 2021.* IEEE, 143–152. https://doi.org/10.34727/2021/isbn.978-3-85448-046-4_23

K. Rustan M. Leino and Peter Müller. 2009. A Basis for Verifying Multi-threaded Programs. In *Programming Languages and Systems, 18th European Symposium on Programming, ESOP 2009, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2009, York, UK, March 22-29, 2009. Proceedings (Lecture Notes in Computer Science, Vol. 5502)*, Giuseppe Castagna (Ed.). Springer, 378–393. https://doi.org/10.1007/978-3-642-00590-9_27

Moritz Lipp, Michael Schwarz, Daniel Gruss, Thomas Prescher, Werner Haas, Anders Fogh, Jann Horn, Stefan Mangard, Paul Kocher, Daniel Genkin, Yuval Yarom, and Mike Hamburg. 2018. Meltdown: Reading Kernel Memory from User Space. In *27th USENIX Security Symposium, USENIX Security 2018, Baltimore, MD, USA, August 15-17, 2018*, William Enck and Adrienne Porter Felt (Eds.). USENIX Association, 973–990. https://www.usenix.org/conference/usenixsecurity18/presentation/lipp

Peter Müller, Malte Schwerhoff, and Alexander J. Summers. 2016. Viper: A Verification Infrastructure for Permission-Based Reasoning. In *Verification, Model Checking, and Abstract Interpretation - 17th International Conference, VMCAI 2016, St. Petersburg, FL, USA, January 17-19, 2016. Proceedings (Lecture Notes in Computer Science, Vol. 9583)*, Barbara Jobstmann and K. Rustan M. Leino (Eds.). Springer, 41–62. https://doi.org/10.1007/978-3-662-49122-5_2

Toby C. Murray, Robert Sison, and Kai Engelhardt. 2018. COVERN: A Logic for Compositional Verification of Information Flow Control. In *2018 IEEE European Symposium on Security and Privacy, EuroS&P 2018, London, United Kingdom, April 24-26, 2018.* IEEE, 16–30. https://doi.org/10.1109/EuroSP.2018.00010

Tobias Nipkow, Lawrence C. Paulson, and Markus Wenzel. 2002. *Isabelle/HOL - A Proof Assistant for Higher-Order Logic.* Lecture Notes in Computer Science, Vol. 2283. Springer. https://doi.org/10.1007/3-540-45949-9

Peter W. O'Hearn. 2004. Resources, Concurrency and Local Reasoning. In *CONCUR 2004 - Concurrency Theory, 15th International Conference, London, UK, August 31 - September 3, 2004, Proceedings (Lecture Notes in Computer Science, Vol. 3170)*, Philippa Gardner and Nobuko Yoshida (Eds.). Springer, 49–67. https://doi.org/10.1007/978-3-540-28644-8_4

Peter W. O'Hearn, Hongseok Yang, and John C. Reynolds. 2004. Separation and Information Hiding. In *Proceedings of the 31st ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages* (Venice, Italy) *(POPL '04)*. Association for Computing Machinery, New York, NY, USA, 268–280. https://doi.org/10.1145/964001.964024

Matthew J. Parkinson and Gavin M. Bierman. 2005. Separation logic and abstraction. In *Proceedings of the 32nd ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL 2005, Long Beach, California, USA, January 12-14, 2005*, Jens Palsberg and Martín Abadi (Eds.). ACM, 247–258. https://doi.org/10.1145/1040305.1040326

Jared Pincus. 2022. *Commutativity Reasoning for the Heap*. Master's thesis. Stevens Institute of Technology.

Alejandro Russo, John Hughes, David A. Naumann, and Andrei Sabelfeld. 2006. Closing Internal Timing Channels by Transformation. In *Advances in Computer Science - ASIAN 2006. Secure Software and Related Issues, 11th Asian Computing Science Conference, Tokyo, Japan, December 6-8, 2006, Revised Selected Papers (Lecture Notes in Computer Science, Vol. 4435)*, Mitsu Okada and Ichiro Satoh (Eds.). Springer, 120–135. https://doi.org/10.1007/978-3-540-77505-8_10

Andrei Sabelfeld and David Sands. 2000. Probabilistic Noninterference for Multi-Threaded Programs. In *Proceedings of the 13th IEEE Computer Security Foundations Workshop, CSFW '00, Cambridge, England, UK, July 3-5, 2000*. IEEE Computer Society, 200–214. https://doi.org/10.1109/CSFW.2000.856937

Daniel Schoepe, Toby Murray, and Andrei Sabelfeld. 2020. VERONICA: Expressive and Precise Concurrent Information Flow Security (Extended Version with Technical Appendices). *CoRR* abs/2001.11142 (2020). arXiv:2001.11142 https://arxiv.org/abs/2001.11142

Geoffrey Smith. 2006. Improved typings for probabilistic noninterference in a multi-threaded language. *J. Comput. Secur.* 14, 6 (2006), 591–623. http://content.iospress.com/articles/journal-of-computer-security/jcs273

Geoffrey Smith. 2007. Principles of Secure Information Flow Analysis. In *Malware Detection*, Mihai Christodorescu, Somesh Jha, Douglas Maughan, Dawn Song, and Cliff Wang (Eds.). Advances in Information Security, Vol. 27. Springer, 291–307. https://doi.org/10.1007/978-0-387-44599-1_13

Geoffrey Smith and Dennis M. Volpano. 1998. Secure Information Flow in a Multi-Threaded Imperative Language. In *POPL '98, Proceedings of the 25th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, San Diego, CA, USA, January 19-21, 1998*, David B. MacQueen and Luca Cardelli (Eds.). ACM, 355–364. https://doi.org/10.1145/268946.268975

Deian Stefan, Pablo Buiras, Edward Z. Yang, Amit Levy, David Terei, Alejandro Russo, and David Mazières. 2013. Eliminating Cache-Based Timing Attacks with Instruction-Based Scheduling. In *Computer Security - ESORICS 2013 - 18th European Symposium on Research in Computer Security, Egham, UK, September 9-13, 2013. Proceedings (Lecture Notes in Computer Science, Vol. 8134)*, Jason Crampton, Sushil Jajodia, and Keith Mayes (Eds.). Springer, 718–735. https://doi.org/10.1007/978-3-642-40203-6_40

Deian Stefan, Alejandro Russo, Pablo Buiras, Amit Levy, John C. Mitchell, and David Mazières. 2012. Addressing covert termination and timing channels in concurrent information flow systems. In *ACM SIGPLAN International Conference on Functional Programming, ICFP'12, Copenhagen, Denmark, September 9-15, 2012*, Peter Thiemann and Robby Bruce Findler (Eds.). ACM, 201–214. https://doi.org/10.1145/2364527.2364557

Viktor Vafeiadis. 2011. Concurrent Separation Logic and Operational Semantics. In *Twenty-seventh Conference on the Mathematical Foundations of Programming Semantics, MFPS 2011, Pittsburgh, PA, USA, May 25-28, 2011 (Electronic Notes in Theoretical Computer Science, Vol. 276)*, Michael W. Mislove and Joël Ouaknine (Eds.). Elsevier, 335–351. https://doi.org/10.1016/j.entcs.2011.09.029

Marco Vassena, Gary Soeller, Peter Amidon, Matthew Chan, John Renner, and Deian Stefan. 2019. Foundations for Parallel Information Flow Control Runtime Systems. In *Principles of Security and Trust - 8th International Conference, POST 2019, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2019, Prague, Czech Republic, April 6-11, 2019, Proceedings (Lecture Notes in Computer Science, Vol. 11426)*, Flemming Nielson and David Sands (Eds.). Springer, 1–28. https://doi.org/10.1007/978-3-030-17138-4_1

Martin T. Vechev, Eran Yahav, Raghavan Raman, and Vivek Sarkar. 2010. Automatic Verification of Determinism for Structured Parallel Programs. In *Static Analysis - 17th International Symposium, SAS 2010, Perpignan, France, September 14-16, 2010. Proceedings (Lecture Notes in Computer Science, Vol. 6337)*, Radhia Cousot and Matthieu Martel (Eds.). Springer, 455–471. https://doi.org/10.1007/978-3-642-15769-1_28

Dennis M. Volpano, Cynthia E. Irvine, and Geoffrey Smith. 1996. A Sound Type System for Secure Flow Analysis. *J. Comput. Secur.* 4, 2/3 (1996), 167–188. https://doi.org/10.3233/JCS-1996-42-304

Dennis M. Volpano and Geoffrey Smith. 1998. Probabilistic Noninterference in a Concurrent Language. In *Proceedings of the 11th IEEE Computer Security Foundations Workshop, Rockport, Massachusetts, USA, June 9-11, 1998*. IEEE Computer Society, 34–43. https://doi.org/10.1109/CSFW.1998.683153

Pengbo Yan and Toby Murray. 2021. SecRSL: security separation logic for C11 release-acquire concurrency. *Proc. ACM Program. Lang.* 5, OOPSLA (2021), 1–26. https://doi.org/10.1145/3485476

Steve Zdancewic and Andrew C. Myers. 2003. Observational Determinism for Concurrent Program Security. In *16th IEEE Computer Security Foundations Workshop (CSFW-16 2003), 30 June - 2 July 2003, Pacific Grove, CA, USA*. IEEE Computer Society, 29. https://doi.org/10.1109/CSFW.2003.1212703