Type-theoretic semantics for transactional concurrency

Aleksandar Nanevski Paul Govereau Greg Morrisett
Harvard University
{aleks,govereau,greg}@eecs.harvard.edu

July 24, 2007

Abstract

We propose a dependent type theory that combines programming, specifications and reasoning about higher-order concurrent programs with shared higher-order transactional memory.

We build on our previous work on Hoare Type Theory (HTT), which is extended here with types that correspond to Hoare-style specifications for transactions. The new types have the form $\mathsf{CMD}\,\{I\}\{P\}\,x{:}A\{Q\}$, and classify concurrent programs that may execute in a shared state with invariant I, and local state precondition P. Upon termination, such programs return a result of type A, and local state changed according to the postcondition Q. Throughout the execution, shared state always satisfies the invariant I, except at specific critical sections which are executed atomically; that is, sequentially, without interference from other processes. Atomic sections may violate the invariant, but must restore it upon exit. We follow the idea of Separation Logic, and adopt "small footprint" specifications, whereby each process is associated with a precondition that tightly describes its state requirement.

HTT is a program logic and a dependently typed programming language at the same time. To consider it as a logic, we prove that it is sound and compositional. To consider it as a programming language, we define its operational semantics, and show adequacy with respect to the specifications.

1 Introduction

Transactional memory is one of the most promising directions in the evolution of concurrent programming languages. It replaces locks, conditional variables, critical regions and other low-level synchronization mechanism, with a higher-level linguistic construct of transactions, and delegates to the run-time system the scheduling of concurrent processes. This frees the programmer from the need to develop potentially complicated and frequently non-modular synchronization protocols that arise in other approaches to concurrency. Transactions make it simpler to write efficient and correct concurrent programs that avoid data races and deadlock. Moreover, transactions are sufficiently well-behaved and compositional to fit naturally into a functional, higher-order language like Haskell [9].

In this paper we are interested not only in programming with transactions, but in developing a formal logic for specification and reasoning about concurrent programs with shared transactional memory. Most program logics for concurrency are versions of Hoare Logic [24]. The recent work on concurrent Separation Logic [22, 3, 29, 6] has made significant inroads into specification and reasoning about shared memory with locking synchronization. The advances of Separation Logic mostly revolve around the idea of spatial separation, whereby each process can be associated with a logical description of exclusive ownership over the state that it requires. This facilitates local reasoning, as the changes that a process makes to its local state do not influence others. Furthermore, Separation Logic connectives lead to particularly convenient descriptions of transferring ownership of state between processes.

When it comes to accessing shared resources, Separation Logic can specify the resource invariants that processes must preserve upon the exclusive use of the resource [22, 3]. Alternatively, Separation Logic can

specify upper and lower bounds on how the shared resource may change, in the style of rely-guarantee reasoning [29, 6].

While Separation Logic has significantly simplified the correctness proofs about shared state, it has mostly concerned itself with imperative *first-order* languages and *low-level synchronization primitives* such as locks. However, irrespective of whether one intends to prove his programs correct or not, programming with such low-level primitives remains difficult. In addition, first-order languages, by definition, do not support advanced linguistic features such as higher-order functions, polymorphism, modules, and abstract data types; all of these are indispensable for programming in the large as they facilitate code reuse, information hiding and modularity. The higher-order abstractions become all the more important if one wants to support specification and reasoning. Yet, most program logics based on (sequential or concurrent) Hoare or Separation Logic have little or no support for these important modularity features.

In this paper we take the step of combining programming, specification, and reasoning in the style of Separation Logic about *higher-order* programs with transactional concurrency. We build on our previous work on Hoare Type Theory (HTT) [20, 19], which is a dependent type theory with extensive support for programming and reasoning about side-effects related to state. Here, we extend HTT with concurrency and transactional primitives.

The main feature of HTT is the *Hoare type*, which takes the form $\mathsf{ST}\,\{P\}\,x{:}A\{Q\}$ and captures partial correctness within the type system. These types classify programs that can be (sequentially) executed in a state satisfying the predicate P and either diverge, or converge to a result $x{:}A$ and a state satisfying Q. In the course of execution, such programs can perform memory reads, writes, allocations and deallocations. By capturing specifications-as-types, HTT makes it possible to abstract over and nest the specifications, combine them with the programs and data that they specify, or even build them algorithmically. All of these features significantly improve over the information hiding and code reuse facilities of Hoare Logic.

From the semantic standpoint, the Hoare type $\mathsf{ST}\,\{P\}\,x{:}A\{Q\}$ is a *monad* [17]. Here we introduce yet another monadic family of Hoare types, which serves to encapsulate concurrent behavior. The new Hoare types take the form $\mathsf{CMD}\,\{I\}\{P\}\,x{:}A\{Q\}$ and classify concurrent programs that execute in a *shared state* satisfying the invariant I, and *local state* satisfying the precondition P. Upon termination, the invariant on the shared state is preserved, but the local state is modified according to the predicate Q. The reader familiar with Haskell's implementation of transactional memory may benefit from the (imprecise) analogy by which Haskell's STM and IO monads correspond to our ST and CMD type families, respectively. For example, as in Haskell, CMD-computations can invoke ST-computations, and fork new CMD threads of computation, but ST-computations are limited to state modifications, in order to facilitate optimistic techniques for implementing transactions.

Similar to Haskell, HTT monads separate the purely functional from the effectful, impure fragment. The pure fragment of HTT includes the Extended Calculus of Constructions (ECC) [12], which is a full dependent type theory with support for abstraction over type universes and predicates in higher-order logic. Indeed, we are currently in the process of implementing HTT in Coq [14], which itself extends ECC with inductive types and predicates (neither of which conflicts with our stateful extensions). For the purposes of this paper, however, we restrict attention to a much smaller fragment which suffices to illustrate our concurrency extensions.

The first technical contribution of this paper, when compared to the previous work, is the formulation of the logical connectives for describing the concurrent behavior. We argue that this logic is sound, and—particularly importantly—compositional. Just as in any type theory, compositionality is expressed by substitution principles, which guarantee that reasoning about HTT programs can be kept local in the sense that the typechecking and verification of a larger program only requires typechecking and verification of its sub-programs, and not any whole-program reasoning.

Just as any type theory, HTT is not only a program logic, but a programming language at the same time[1]. As the second contribution of the current paper, we endow the stateful and concurrent terms of HTT with operational semantics, and prove that this operational semantics is adequate for the intended interpretation of the Hoare types.

[1]Hence exhibiting a variation on the Curry-Howard isomorphism.

The rest of the paper is structured as follows. In Section 2 we introduce the basic stateful and transactional constructs, and illustrate how programs can be specified using Hoare types. In Section 3 we describe the formal syntax of the language, the connection with some well-known features from Hoare Logic, like ghost variables, and the definitions of the relational connectives that will serve to capture the semantics of state and concurrency. Section 4 presents the type system, and Section 5 describes the basic theorems about it. In Section 6 we introduce the operational semantics, and the proof of its adequacy. Section 7 discusses the related and future work, and Section 8 concludes.

2 Overview of monadic state and transactional memory

There are three conceptual levels in HTT: the purely functional fragment, the ST fragment, and the CMD fragment. As the name suggests, the pure fragment has no computational effects. The ST fragment includes sequential stateful commands $\mathsf{alloc}\, M$ (allocation), $!\, M$ (read), $M_1{:=}M_2$ (write), and $\mathsf{dealloc}\, M$ (deallocation). In addition, the ST fragment contains conditionals, and allows one to construct recursive (i.e., possibly diverging) computations. The CMD fragment includes commands $\mathsf{atomic}\ E$ (atomically run the ST-computation E), and $x_1{\Leftarrow}E_1 \| x_2{\Leftarrow}E_2$ (run the two CMD computations E_1 and E_2 in parallel). The CMD fragment also includes a publish primitive which will be explained below, as well as constructors for conditionals and recursion.

The stateful sequential computations are classified by types of the form $\mathsf{ST}\,\{P\}\,x{:}A\{Q\}$, where P and Q are pre- and postconditions on the state. To illustrate these types, and their interaction with lambda abstraction and function types from the pure fragment, consider the function incBy, which takes a pointer l to a nat, a value n:nat, and then increments the contents of l by n. This function can be implemented as follows.

$$
\begin{array}{rl}
\mathsf{incBy} \ : & \Pi l{:}\mathsf{loc}.\, \Pi n{:}\mathsf{nat}. \\
& \quad [v{:}\mathsf{nat}].\, \mathsf{ST}\,\{l \mapsto_{\mathsf{nat}} v\}\, x{:}1\{l \mapsto_{\mathsf{nat}} v + n\} \\
= & \lambda l.\, \lambda n.\, \mathsf{stdo}\ (t \Leftarrow\, !\, l; \\
& \qquad\qquad\quad l := t + n; \\
& \qquad\qquad\quad \mathsf{return}\ (\,))
\end{array}
$$

The term syntax is chosen to closely resemble Haskell's do-notation, but also to support the meta-theoretic development (i.e., substitution principles). The keyword stdo encapsulates in its scope the stateful part of the code, separating it from the functional abstraction. The stateful code first reads from the location l and binds the obtained value to the (immutable) temporary variable t ($t \Leftarrow\, !\, l$), then writes back the increased value ($l := t + n$), before returning $(\,){:}1$.

The type of incBy is a bit more involved: It specifies that incBy takes two arguments l:loc and n:nat, and returns a block of stateful code with Hoare type $[v{:}\mathsf{nat}].\, \mathsf{ST}\,\{l \mapsto_{\mathsf{nat}} v\}\, x{:}1\{l \mapsto_{\mathsf{nat}} v + n\}$. As expected, the precondition $l \mapsto_{\mathsf{nat}} v$ requires that at the beginning of the stateful block, the location l *points to* some value v:nat. The postcondition ensures that at the end l points to an incremented value. We note the use of the variable v:nat, which is bound in the Hoare type, with the scope extending through the precondition and the postcondition. The variable v serves to relate the value stored in the initial state of the monad, with the value at the ending state. In accordance with the standard Hoare logic terminology, we call v a *ghost variable*.

Concurrent computations are classified by types of the form $\mathsf{CMD}\,\{I\}\{P\}\,x{:}A\{Q\}$, where P and Q are pre- and post-conditions on the local state of the computation, and I is an invariant on the state that is shared with other processes. The key construct mediating access to shared state is the atomic primitive. It presents the programmer with the abstraction that the enclosed block of code executes sequentially, and in isolation from all the other parallel processes. Of course, implementations are not so naive and extract parallelism through advanced run-time techniques. In particular, atomic blocks are optimistically run in parallel with the hope that the blocks will not perform conflicting changes to memory. To handle the case where there is a conflict, the runtime system aborts one of the conflicting blocks by rolling back its changes to the store and then re-starting the block.

Conceptually, the atomic primitive has the following type (although the exact details in Section 4 will differ somewhat, to account for the technical requirements of the type system):

$$\mathsf{atomic} : \mathsf{ST}\,\{I * P\}\,x{:}A\{I * Q\} \to \mathsf{CMD}\,\{I\}\{P\}\,x{:}A\{Q\}$$

We can think of a thread running the command $\mathsf{atomic}\,M$, as acquiring a global lock on the shared state, executing the sequential code M, and then releasing the lock. During the atomic block, the thread is allowed to access both global and local state. Upon entry to the block, the global state is described by the invariant I, and the local state is described by P. Furthermore, we are guaranteed that the local and global state are disjoint through the use of the *separating conjunction* specification $I * P$. Throughout the execution of the atomic block, the thread is allowed to read and modify both the local and global state described by the specification. In particular, it can safely violate the invariant on the global state since no other thread can see the changes during the transaction. Furthermore, the thread is able to freely transfer locations from the local state to the global state and vice versa. Upon termination of the block, the thread must re-establish that the heap can be split into a local portion, now described by Q, and a global portion once again described by the invariant I, resulting in a post-condition of $I * Q$. In summary, a sequential command with type $\mathsf{ST}\,\{I{*}P\}\,x{:}A\{I * Q\}$ can be lifted via atomic to a concurrent command with interface $\mathsf{CMD}\,\{I\}\{P\}\,x{:}A\{Q\}$.

As a simple example, consider the following definition:

$$\begin{array}{l}\mathsf{transfer} = \lambda l_1, l_2, n.\,\mathsf{cmdo}\\ \quad (t \Leftarrow \mathsf{atomic}(\\ \qquad t_1 \Leftarrow !\,l_1;\\ \qquad \mathsf{if}\ t_1 < n\ \mathsf{then\ return\ ff}\\ \qquad \mathsf{else}\ (\mathsf{decBy}\ l_1\ n;\ \mathsf{incBy}\ l_2\ n;\ \mathsf{return\ tt}));\\ \quad \mathsf{return}\ t)\end{array}$$

The transfer command attempts to atomically transfer the value n from location l_1 to location l_2, using the auxiliary commands incBy and decBy (not shown here). If l_1 holds a value less than n, then the transfer aborts and returns boolean ff, but if the transfer is successful, the command returns the boolean tt.

We can assign transfer a number of types, depending upon what correctness properties we wish to enforce. For example, in a banking application, we may wish to capture the constraint that the sum of the balances of the accounts must remain constant. That is, money can only be transferred, but not created or destroyed. In such a setting, we can use the following type:

$$\begin{array}{rcl}\mathsf{transfer} & : & \Pi l_1{:}\mathsf{loc}.\,\Pi l_2{:}\mathsf{loc}.\,\Pi n{:}\mathsf{nat}.\\ & & \quad \mathsf{CMD}\,\{I(l_1,l_2)\}\{\mathsf{emp}\}\,x{:}\mathsf{bool}\{\mathsf{emp}\}\end{array}$$

where $I(l_1, l_2) = \exists v_1{:}\mathsf{nat}.\,\exists v_2{:}\mathsf{nat}.\,((l_1 \mapsto_{\mathsf{nat}} v_1) * (l_2 \mapsto_{\mathsf{nat}} v_2)) \wedge (v_1 + v_2 = k)$. Here, emp denotes an empty store, and $l \mapsto_\tau v$ denotes a store where location l points to a value v of type τ. Thus, the specification of transfer captures the invariant that the sum of the values in l_1 and l_2 must equal the constant k. Note that during the transfer, the invariant is violated, but is eventually restored. Thus, irrespective of the number of transfers executed between l_1 and l_2, the sum of the values stored into these locations always remains k. Note also that transfer operates only on shared state, and imposes no requirements on the local state. In particular, it can run even if the local state is empty, and any extensions of the local state will not be touched. In HTT, like in Separation Logic, this property is specified by using the predicate emp as a precondition, to tightly describe the local space requirements of the function.

We can now execute a number of transfers between l_1 and l_2 concurrently; the system will take care to preserve the invariant.

$$\begin{array}{rcl}\mathsf{transfer}_2 & : & \Pi l_1{:}\mathsf{loc}.\,\Pi l_2{:}\mathsf{loc}.\\ & & \quad \mathsf{CMD}\,\{I(l_1,l_2)\}\{\mathsf{emp}\}\,x{:}\mathsf{bool}\{\mathsf{emp}\}\\ & = & \mathsf{cmdo}((t_1 \Leftarrow \mathsf{transfer}\ l_1\ l_2\ 10 \parallel\\ & & \quad t_2 \Leftarrow \mathsf{transfer}\ l_2\ l_1\ 20);\\ & & \quad \mathsf{return}(t_1\ \mathrm{and}\ t_2))\end{array}$$

The above function forks two processes to concurrently execute two transfers, one between l_1 and l_2 and the other between l_2 and l_1. The values obtained as a result of each process are collected into variables t_1 and t_2, and the function returns tt if both transfers succeed.

2.1 Guarded commands

As a more interesting example, we next develop a function guard which waits in a busy loop until a provided location contains some required value[2]. The guard definition will be a function of four arguments, so that guard $\alpha\ l\ n\ f$ reads the contents of location l, and loops until this contents equals n. Then it will execute the ST command f atomically, and return the obtained value of type α. For example, guard $1\ l_1\ 42$ (decBy $l_1 35$) will wait until l_1 contains 42, and then decrement its contents by 35.

$$\begin{array}{l}\mathsf{guard} : \forall\alpha.\,\Pi l{:}\mathsf{loc}.\,\Pi n{:}\mathsf{nat}.\\ \quad \mathsf{ST}\,\{(l \mapsto_{\mathsf{nat}} n) * J * P\}\,x{:}\alpha\{(l \mapsto_{\mathsf{nat}} -) * J * Q(x)\} \to\\ \quad\quad \mathsf{CMD}\,\{(l \mapsto_{\mathsf{nat}} -) * J\}\{P\}\,x{:}\alpha\{Q(x)\}\end{array}$$

The return type of guard is a CMD-monad in order to allow other processes to concurrently set the value of l, while guard is busy waiting. Correspondingly, l should be a shared location, requiring the shared state invariant of the CMD-type to specify that $l \mapsto_{\mathsf{nat}} -$. Whatever the precondition P and postcondition Q on the local state this return type has, the ST-computation that is executed atomically should augment them with the knowledge that l is allocated, and that l contains value n at the beginning of the atomic execution. We further allow that the shared state may include an additional section described by the predicate J. This section can be modified by the ST-computation, as long as the validity of J is preserved. Notice that guard is implicitly polymorphic in the predicates J, P and Q. In this paper we do not discuss explicit polymorphism over predicates, but notice that such a feature is available in ECC and Coq, and we have already shown in the previous work that HTT can consistently (and usefully) be extended with it [19].

We split the implementation of guard into two parts. We first assume a helper function waitThen which carries out the busy loop, but instead of immediately returning the result of the atomic execution, it stores this result into a temporary location r. Using waitThen, guard is implemented as follows.

$$\begin{array}{l}\mathsf{guard} = \Lambda\alpha.\,\lambda l.\,\lambda n.\,\lambda st.\\ \quad \mathsf{cmdo}\ (r \Leftarrow \mathsf{atomic}(t \Leftarrow \mathsf{alloc}\,0; \mathsf{return}\ t);\\ \quad\quad \mathsf{waitThen}\ \alpha\ l\ n\ st\ r;\\ \quad\quad t \Leftarrow \mathsf{atomic}(x \Leftarrow !\,r; \mathsf{dealloc}\ r; \mathsf{return}\ x);\\ \quad\quad \mathsf{return}\ t)\end{array}$$

The code first allocates the temporary location r, then waits on l, expecting the result of waiting to show up in r. Finally, it reads the result from r, and passes it out but only after r is deallocated. Notice that the accesses to store are always within an atomic block.

$$\begin{array}{l}\mathsf{waitThen} = \Lambda\alpha.\,\lambda l.\,\lambda n.\,\lambda st.\,\lambda r.\\ \quad \mathsf{cmdo}(t \Leftarrow \mathsf{fix}(\lambda c.\,\mathsf{cmdo}\\ \quad\quad (ok \Leftarrow \mathsf{atomic}(x \Leftarrow !\,l;\\ \quad\quad\quad \mathsf{if}\,(x = n)\,\mathsf{then}\\ \quad\quad\quad\quad y \Leftarrow st;\\ \quad\quad\quad\quad r := y;\\ \quad\quad\quad\quad \mathsf{return}\ \mathsf{ff}\\ \quad\quad\quad \mathsf{else\ return\ tt});\\ \quad\quad \mathsf{if}\ ok\ \mathsf{then}\ x \Leftarrow c; \mathsf{return}\ x\\ \quad\quad \mathsf{else\ return}\ ()));\\ \quad\quad \mathsf{return}\ t)\end{array}$$

[2]Of course, a real implementation will provide something like guard as a blocking primitive instead of encoding it via busy-waiting.

Under the fixpoint, waitThen first atomically reads l, and based on the value, either executes st (by the command $y \Leftarrow st$), storing the result y into r, or simply exits the atomic block. Either way, it passes back via the flag ok the information about which branch was taken. If the contents of l was not appropriate, it goes around the loop again, by invoking the fixpoint computation c. Otherwise, r must contain the required value, so the function exits. The type of waitThen is

$$\begin{array}{l}\mathsf{waitThen} : \forall\alpha.\,\Pi l{:}\mathsf{loc}.\,\Pi n{:}\mathsf{nat}.\,\Pi r{:}\mathsf{loc}.\\ \qquad \mathsf{ST}\{(l \mapsto_{\mathsf{nat}} n) * J * P\}x{:}\alpha\{(l \mapsto_{\mathsf{nat}} -) * J * Q(x)\}\\ \qquad\quad \to \mathsf{CMD}\{(l \mapsto_{\mathsf{nat}} -) * J\}\{P * (r \mapsto_{\mathsf{nat}} 0)\}\,t{:}1\\ \qquad\qquad \{\exists x{:}\alpha.\,Q(x) * (r \mapsto_{\alpha} x)\}\end{array}$$

Notice that the return CMD type requires the existence of the location r in the local state, and guarantees that r contains the result of the atomic execution at the end. The later is ensured by the spatial conjunction with $Q(x)$ in the postcondition. Also, the first two components of the return type represents the loop invariant of the busy loop of waitThen. Essentially, throughout the iterations, we know that $(l \mapsto_{\mathsf{nat}} -) * J$ holds for the shared store and $P * (r \mapsto_{\mathsf{nat}} 0)$ holds for the local store.

2.2 Synchronizing variables

Using guard, we can now implement synchronizing variables (also know as "MVars" in Haskell). A synchronizing variable is a location in memory that either contains a value, or is empty, with two operations: put and take. The put operation will put a new value into an empty variable, and block otherwise. The take operation will block until a variable becomes full, then read the value from a full variable, emptying it.

We implement each synchronizing variable using two locations l and v in the shared heap, l for the empty/full flag, and v for the value. The invariant for the shared heap is $I_{sv}(l,v) = \exists n{:}\mathsf{nat}.\,((l \mapsto_{\mathsf{nat}} n) \wedge (n = 0 \vee n = 1)) * (v \mapsto_A -)$, requiring that l points to a nat (0 for empty, 1 for full), and that v contains a value of a fixed type. Both put and take are CMD-computations over a shared heap described by I_{sv}. Since these operations only operate on the shared state, the pre- and postconditions on the local state are trivially emp. The implementations call the guard function, instantiated with $J = (v \mapsto_A -)$, and $P = Q = \mathsf{emp}$.

$$\begin{array}{rl}\mathsf{put} & : \Pi l{:}\mathsf{loc}.\,\Pi v{:}\mathsf{loc}.\,A \to \mathsf{CMD}\,\{I_{sv}(l,v)\}\{\mathsf{emp}\}\,x{:}1\{\mathsf{emp}\}\\ & = \lambda l.\,\lambda v.\,\lambda x.\,\mathsf{guard}\ 1\ l\ 0\ \mathsf{stdo}\ (l := 1; v := x; \mathsf{return}\ ())\end{array}$$

$$\begin{array}{rl}\mathsf{take} & : \Pi l{:}\mathsf{loc}.\,\Pi v{:}\mathsf{loc}.\,\mathsf{CMD}\,\{I_{sv}(l,v)\}\{\mathsf{emp}\}\,x{:}A\{\mathsf{emp}\}\\ & = \lambda l.\,\lambda v.\,\mathsf{guard}\ A\ l\ 1\ \mathsf{stdo}\ (l := 0; x \Leftarrow !\,v; \mathsf{return}\ x)\end{array}$$

We can test for fullness/emptiness, without blocking, by using the function empty, similar to the one provided in Haskell standard libraries [9].

$$\begin{array}{rl}\mathsf{empty} & : \Pi l{:}\mathsf{loc}.\,\Pi v{:}\mathsf{loc}.\,\mathsf{CMD}\,\{I_{sv}(l,v)\}\{\mathsf{emp}\}\,x{:}\mathsf{nat}\{\mathsf{emp}\}\\ & = \lambda l.\,\lambda v.\,\mathsf{cmdo}(t \Leftarrow \mathsf{atomic}(x \Leftarrow !\,l; \mathsf{return}\ x);\\ & \qquad\qquad \mathsf{return}\ t)\end{array}$$

2.3 Producer-consumer pattern

HTT includes an additional concurrency primitive publish J, which logically takes a part of the local state described by the predicate J and moves it into the shared state. A computation may need to perform this operation if it wants to spawn some child processes to execute concurrently on the given local state. We illustrate the primitive by building a producer-consumer pattern, whereby we allocate a new synchronizing variable, publish it as shared state, and launch two processes which communicate via the now shared variable. The shared variable becomes a primitive communication channel between the processes.

Suppose that we have a producer function, p, and a consumer function, c. Then, we can easily construct functions which read from and write to a shared variable using p and c.

$$\begin{aligned}\mathsf{produce} = \lambda l.\, \lambda v.\, \mathsf{cmdo}(t = \mathsf{fix}\ \lambda f.\, \mathsf{cmdo}(x \Leftarrow p; \mathsf{put}\ l\ v\ x;\\ s \Leftarrow f; \mathsf{return}\ s);\\ q \Leftarrow t; \mathsf{return}\ q)\end{aligned}$$

$$\begin{aligned}\mathsf{consume} = \lambda l.\, \lambda v.\, \mathsf{cmdo}(t = \mathsf{fix}\ \lambda f.\, \mathsf{cmdo}(x \Leftarrow \mathsf{take}\ l\ v; c\ x;\\ s \Leftarrow f; \mathsf{return}\ s);\\ q \Leftarrow t; \mathsf{return}\ q)\end{aligned}$$

Here, p and c both obtain a CMD-computations with a shared invariant $I_{sv}(l, v)$. In order to use produce and consume, we must first establish this invariant; this is where publish comes in.

$$\begin{aligned}&\mathsf{cmdo}(l \Leftarrow \mathsf{atomic}(t \Leftarrow \mathsf{alloc}\ 0; \mathsf{return}\ t);\\ &\qquad v \Leftarrow \mathsf{atomic}(t \Leftarrow \mathsf{alloc}\ a; \mathsf{return}\ t);\\ &\qquad \mathsf{publish}(I_{sv}(l, v));\\ &\qquad x_1 \Leftarrow \mathsf{produce}\, l\, v \parallel x_2 \Leftarrow \mathsf{consume}\, l\, v;\\ &\qquad \mathsf{return}\ (\,))\end{aligned}$$

After allocating l and v, we publish them with the invariant I_{sv}. After the publish, produce and consume can execute in parallel, and each has access to l and v as shared state.

3 Formal syntax and definitions

In this section we present the syntax of HTT, and discuss the constructs in more detail.

Types	A, B, τ	::=	$\alpha \mid \mathsf{bool} \mid \mathsf{nat} \mid 1 \mid \Pi x{:}A.\, B \mid \forall \alpha.\, A \mid \mathsf{ST}\,\{P\}\,x{:}A\{Q\} \mid \mathsf{CMD}\,\{I\}\{P\}\,x{:}A\{Q\}$
Predicates	P, Q, R, I	::=	$\mathsf{id}_A(M, N) \mid \mathsf{seleq}_\tau(H, M, N) \mid \top \mid \bot \mid P \wedge Q \mid P \vee Q \mid P \supset Q \mid \neg P \mid$ $\forall x{:}A.\, P \mid \forall \alpha.\, P \mid \forall h{:}\mathsf{heap}.\, P \mid \exists x{:}A.\, P \mid \exists \alpha.\, P \mid \exists h{:}\mathsf{heap}.\, P$
Heaps	H, G	::=	$h \mid \mathsf{empty} \mid \mathsf{upd}_\tau(H, M, N)$
Terms	K, L, M, N	::=	$x \mid \mathsf{tt} \mid \mathsf{ff} \mid \bar{n} \mid M \oplus N \mid (\,) \mid \lambda x.\, M \mid K\ M \mid \Lambda \alpha.\, M \mid K\ \tau \mid$ $\mathsf{stdo}\ E \mid \mathsf{cmdo}\ E \mid M : A$
Computations	E, F	::=	$\mathsf{return}\ M \mid x \Leftarrow K; E \mid x \Leftarrow !_\tau\, M; E \mid M :=_\tau N; E \mid x \Leftarrow \mathsf{alloc}_\tau\, M; E \mid$ $\mathsf{dealloc}\, M; E \mid x \Leftarrow \mathsf{atomic}_{A,P,Q}\, E_1; E \mid$ $(x_1{:}A_1 \Leftarrow E_1{:}P_1 \parallel x_2{:}A_2 \Leftarrow E_2{:}P_2); E \mid \mathsf{publish}\ I; E \mid x \Leftarrow \mathsf{fix}_A\, M; E \mid$ $x \Leftarrow \mathsf{if}_A\ M\ \mathsf{then}\ E_1\ \mathsf{else}\ E_2; E$
Contexts	Δ	::=	$\cdot \mid \Delta, x{:}A \mid \Delta, \alpha \mid \Delta, h{:}\mathsf{heap} \mid \Delta, P$

3.1 Types

In addition to the already described Hoare types, HTT admits the types of booleans and natural numbers, dependent function types $\Pi x{:}A.\, B$, and polymorphic quantification $\forall \alpha.\, A$. The type variables α in polymorphic quantification ranges over monomorphic types only, as customary in, say, Standard ML (SML). Thus, HTT supports only predicative polymorphism, although extensions with impredicativity are currently being investigated by Petersen et al. [25]. As usual with dependent types, we write $A \to B$ instead of $\Pi x{:}A.\, B$, when the type B does not depend on x. In the current paper, we ignore the other useful type constructors from pure type theories, like Σ-types and inductive types. These do not present any theoretical problems. For example, we have studied the extension with Σ-types in the previous work [19], and have also implemented the stateful (but concurrency-free) part of HTT in Coq, which supports inductive types and predicates. These extensions, however, do add bulk to the development, so we omit them here in order to focus on the main ideas related to concurrency.

3.2 Terms

The purely functional fragment consists of the usual term constructors: boolean values, numerals $\bar{n}$ and the basic arithmetic operations (collectively abbreviated as $M \oplus N$), the unit value $(\,):1$, lambda abstraction and application, and type abstraction and application. We do not annotate lambda abstractions with the domain types, but instead provide a constructor $M{:}A$ to ascribe a type A to the term M. This organization facilitates bidirectional typechecking [30]. The stdo and cmdo constructors are the introduction forms for the corresponding Hoare types, analogous to the monadic-do in Haskell, except in HTT we have separate constructor for each monad, to avoid any confusion.

3.3 Computations

The scope of stdo and cmdo is a computation, which is a semi-colon separated list of commands, terminating with a return value. We have already described the intuition behind most of the constructors in Section 2. However, some of these require explicit annotations with types, pre/postconditions and invariants, which were omitted before, so we now revisit them with the additional details.

For example, HTT supports strong updates by which a location pointing to a value of type τ_1 may be updated with a value of some other type τ_2. Correspondingly, the ST primitives for reading, writing and allocation must be annotated with the type of the manipulated value.

CMD-computations are annotated as follows. (1) $(x_1{:}A_1 \Leftarrow E_1{:}P_1 \parallel x_2{:}A_2 \Leftarrow E_2{:}P_2)$ *forks* two parallel child processes E_1 and E_2. Upon their termination, the processes are *joined*, that is, their return results are bound to x_1 and x_2, respectively, their private space is returned to the parent process, and the execution proceeds to the subsequent command of the parent process. The $\parallel$ command explicitly requires the return types A_1 and A_2 and the preconditions P_1 and P_2 on the parallel processes. The preconditions indicate which part of the local state of the parent process is delegated to each. The split of the local state must be disjoint. If the two processes are supposed to share state, then that state must be declared as shared. (2) $x \Leftarrow \mathsf{atomic}_{A,P,Q}\, E_1$ explicitly requires the return type of E_1 as well as the precondition P and postcondition Q on the local state that E_1 manipulates. This local state will be joined with the shared state of the parent process, before E_1 executes atomically. (3) publish I does not require additional annotation beyond the predicate I. However, we mention here that I must be *precise*, in the sense that it uniquely defines the heap fragment that should be published. For example, the predicate $(x \mapsto_{\mathsf{nat}} - * y \mapsto_{\mathsf{nat}} -)$ is precise, and would correspond to publishing the locations x and y. On the other hand, the predicate $\top$ is not precise, as it holds of every heap. Precision is customarily required of shared state invariants in Separation Logic [3, 22].

Finally, the conditional and the fix constructs are present in both monads. The conditional is annotated with the expected types of its branches. Fix is annotated with the type A, and computes the least fixed point of the function $M{:}A{\to}A$. Here A must be a Hoare type, in order to guarantee that all uses of recursion (and hence, potential occasions for non-termination) appear under the guard of stdo or cmdo. Thus, non-termination in HTT is considered an effect, and is encapsulated under the monad, in order to preserve the logical properties of the underlying pure calculus. In particular, the encapsulation would prevent the recursion from unrolling in the reductions performed by equational reasoning during typechecking.

3.4 Heaps

In HTT we model heap locations by natural numbers, although in the examples we write loc instead of nat to emphasize when a natural number is used as a pointer. Heaps are modeled as functions mapping a location N to a pair (τ, M) where $M{:}\tau$. In this case, we say that N *points to* M, or that M is the *contents* of N. The type τ is required to be monomorphic, in order to preserve the predicativity of the type theory. This is analogous to the treatment of state in, for example, SML. However, τ can be a dependent function type, as well as a Hoare type. Thus, heaps in HTT are *higher-order*, albeit predicative.

Syntactically, we build heaps out of the following primitives: (1) empty stands for the empty heap, that is, a nowhere defined function. (2) $\mathsf{upd}_\tau(H, M, N)$ is a function which returns $N{:}\tau$ at argument M, but equals H at other arguments. It models the heap obtained from H by writing $N{:}\tau$ into the address M.

3.5 Predicate logic and heap semantics

In this paper we consider a first-order, polymorphic, predicate logic over heaps. We have shown in the previous work [19] that the logic can easily be scaled to higher-order, simply by introducing a new type of propositions (as customary in ECC or Coq). The restriction to first-order will suffice for the current paper.

Aside from the usual connectives of first-order logic, we provide two primitives: (1) $\mathsf{id}_A(M,N)$ is the equality at type A. We will frequently abbreviate it as $M =_A N$ or simply $M = N$. (2) $\mathsf{seleq}_\tau(H,M,N)$ reflects the semantics of heaps into the assertion logic of the Hoare types. It holds iff the heap H contains $N{:}\tau$ at location M. The following axioms relate seleq and update.

$$\begin{array}{l}\neg\mathsf{seleq}_\tau(\mathsf{empty},M,N)\\ \mathsf{seleq}_\tau(\mathsf{upd}_\tau(H,M,N),M,N)\\ M_1\neq M_2\wedge\mathsf{seleq}_\tau(\mathsf{upd}_\sigma(H,M_1,N_1),M_2,N_2)\supset\mathsf{seleq}_\tau(H,M_2,N_2)\\ \mathsf{seleq}_\tau(H,M,N_1)\wedge\mathsf{seleq}_\tau(H,M,N_2)\supset N_1=_\tau N_2\end{array}$$

The first axiom states that an empty heap does not contain any assignments. The second and the third are the well-known McCarthy axioms for functional arrays [16]. The fourth axiom asserts a version of heap functionality: a heap may assign at most one value to a location, for each given type. The fourth axiom is slightly weaker than expected, as we would like to state that a heap assigns at most one type and value to a location. This is easily expressible in the extension of HTT with higher-order logic [19].

3.6 Separation logic

Given the heaps as above, we can now define predicates expressing heap equality, disjointness, and disjoint union of heaps [20].

$$\begin{array}{rcl}P\subset\!\supset Q&=&P\supset Q\wedge Q\supset P\\ H_1=H_2&=&\forall\alpha.\forall x{:}\mathsf{nat}.\forall v{:}\alpha.\,\mathsf{seleq}_\alpha(H_1,x,v)\subset\!\supset\mathsf{seleq}_\alpha(H_2,x,v)\\ M\in H&=&\exists\alpha.\exists v{:}\alpha.\,\mathsf{seleq}_\alpha(H,M,v)\\ M\notin H&=&\neg(M\in H)\\ \mathsf{share}(H_1,H_2,M)&=&\forall\alpha.\forall v{:}\alpha.\;\;\mathsf{seleq}_\alpha(H_1,M,v)\subset\!\supset\mathsf{seleq}_\alpha(H_2,M,v)\\ \mathsf{splits}(H,H_1,H_2)&=&\forall x{:}\mathsf{nat}.\,(x\notin H_1\wedge\mathsf{share}(H,H_2,x))\vee(x\notin H_2\wedge\mathsf{share}(H,H_1,x))\end{array}$$

In English, $\supset$ is logical implication, $\subset\!\supset$ is logical equivalence, $H_1 = H_2$ is heap equality, $M \in H$ iff the heap H assigns to the location M, share states that H_1 and H_2 agree on the location M, and splits states that H can be split into disjoint heaps H_1 and H_2.

We next formally define the assertions familiar from Separation Logic [21]. All of these are relative to the free variable m, which denotes the current heap fragment of reference. We will call predicates with one free heap variable m *unary predicates*, and use letters P, R, S and I to range over them. Given a unary predicate P, we will customarily use the syntax for functional application, and write $P\ H$ as an abbreviation for $[H/\mathsf{m}]P$.

$$\begin{array}{lcl}\mathsf{emp}&=&\mathsf{m}=\mathsf{empty}\\ M\mapsto_\tau N&=&\mathsf{m}=\mathsf{upd}_\tau(\mathsf{empty},M,N))\\ M\hookrightarrow_\tau N&=&\mathsf{seleq}_\tau(\mathsf{m},M,N)\\ P*S&=&\exists h_1{:}\mathsf{heap}.\exists h_2{:}\mathsf{heap}.\,\mathsf{splits}(\mathsf{m},h_1,h_2)\wedge P\ h_1\wedge S\ h_2\\ P-\!\!*S&=&\forall h_1{:}\mathsf{heap}.\forall h_2{:}\mathsf{heap}.\,\mathsf{splits}(h_2,h_1,\mathsf{m})\supset P\ h_1\supset S\ h_2\\ \mathsf{this}\,H&=&\mathsf{m}=H\\ \mathsf{precise}\ P&=&\forall h_1,h_1',h_2,h_2'{:}\mathsf{heap}.\,\mathsf{splits}(\mathsf{m},h_1,h_1')\supset\mathsf{splits}(\mathsf{m},h_2,h_2')\supset P\ h_1\supset P\ h_2\supset h_1=h_2\end{array}$$

We have already given the informal descriptions of emp, $M\mapsto_\tau N$ and $P*S$ in Section 2. $M\hookrightarrow_\tau N$ iff current heap contains *at least* the location M pointing to $N{:}\tau$. $P-\!\!*S$ holds iff any extension of the current heap by a heap satisfying P, produces a heap satisfying S. this (H) iff the current heap equals H. Concerning the last predicate, precise P holds iff for any given heap m, there is *at most one* subheap h such that $P\ h$.

With these definitions, it should now be apparent that the preconditions, postconditions and invariants in our Hoare types are predicates over heaps, and that they implicitly depend on the heap variable m. For example, the type $\mathsf{ST}\,\{\mathsf{emp}\}\,x{:}A\{\mathsf{emp}\}$ really equals $\mathsf{ST}\,\{\mathsf{m} = \mathsf{empty}\}\,x{:}A\{\mathsf{m} = \mathsf{empty}\}$, where m in the precondition denotes the initial heap, and m in the postcondition denotes the ending heap of a stateful computation. Thus, the two m variables are really different. If we wanted to make the scope of m explicit, we would write the type $\mathsf{ST}\,\{P\}\,x{:}A\{Q\}$ explicitly as

$$\mathsf{ST}\,\{\mathsf{m}.\,P\}\,x{:}A\{\mathsf{m}.\,Q\}$$

However, in order to reduce clutter, we leave the bindings of m implicit. We adopt a similar strategy for $\mathsf{CMD}\,\{I\}\{P\}\,x{:}A\{Q\}$, where I also depends on an implicitly bound variable m.

3.7 Ghost variables and binary postconditions

The programs from Section 2 already exhibit that the use of Hoare types frequently requires a set of ghost variables that scope over the precondition and the postcondition in order to relate the two. For example, the program incBy with the type

$$\mathsf{incBy} : \Pi l{:}\mathsf{loc}.\,\Pi n{:}\mathsf{nat}.\,[v{:}\mathsf{nat}].\,\mathsf{ST}\,\{l \mapsto_{\mathsf{nat}} v\}\,x{:}1\{l \mapsto_{\mathsf{nat}} v + n\}$$

needs the ghost variable v to name the value initially stored into l.

Unfortunately ghost variables are inconvenient for the semantics of HTT, for several reasons. For one, a binding construct like the [brackets] above complicates the definition of equality between Hoare types: is the context with a ghost variable $x{:}A{\times}B$ equal to the context with two ghost variables $x_1{:}A, x_2{:}B$? Another problem arises if one wants to consider abstraction over predicates, as is frequently necessary in practical applications in order to hide the invariants of the local state of functions and modules [19]. Then we quickly face the need to quantify over contexts of ghost variables.

To avoid dealing with these issues, we need an alternative to ghost variables, which would still let us relate the preconditions and postconditions of Hoare types. One type-theoretic possibility is to explicitly functionally abstract the ghost variables of each Hoare type. For example, incBy may be re-typed as:

$$\begin{aligned}\mathsf{incBy}' \quad : \quad & \Pi l{:}\mathsf{loc}.\,\Pi n{:}\mathsf{nat}.\,\Pi v{:}\mathsf{nat}.\\ & \mathsf{ST}\,\{l \mapsto_{\mathsf{nat}} v\}\,x{:}1\{l \mapsto_{\mathsf{nat}} v + n\}\end{aligned}$$

This, however, is not a convenient solution either. Functional abstraction will require every caller of incBy' to instantiate v at run-time. The ghost variable v, which should serve only the purpose of logically connecting the precondition and the postcondition of the Hoare type, suddenly acquires a computational significance; it has to be explicitly supplied by the caller, and the value, when instantiated, has to produce a precondition that is true at the given program point. More concretely, in order to increment the contents of l by executing incBy', the caller must already know what the value stored in l is. This, of course, makes the usefulness of incBy' quite dubious. If the caller already knows the stored value, why not simply write its increment back into l directly?

A better alternative, and the one that we adopt here, is to allow that postconditions not only depend on the variable m denoting the current heap at the end of the computation, but also on the variable i that denotes the initial heap. That is, if we made the scopes explicit, then the type $\mathsf{ST}\,\{P\}\,x{:}A\{Q\}$ would be written as $\mathsf{ST}\,\{\mathsf{m}.\,P\}\,x{:}A\{\mathsf{i}.\,\mathsf{m}.\,Q\}$. The second heap variable in the postcondition can be used to relate the values stored in the initial heap, to the values stored in the ending heap. The type of incBy may be written as

$$\begin{aligned}\mathsf{incBy}'' : \Pi l{:}\mathsf{loc}.\,\Pi n{:}\mathsf{nat}.\,\mathsf{ST}\{\exists v.\,l \mapsto_{\mathsf{nat}} v\}\ r : 1\\ \{\forall v.\,(l \mapsto_{\mathsf{nat}} v)\ \mathsf{i} \supset (l \mapsto_{\mathsf{nat}} v)\ \mathsf{m}\}\end{aligned}$$

Under this binding convention, the syntax of Hoare types with ghost variables becomes just a syntactic sugar. The Hoare type $[\Delta].\,\mathsf{ST}\,\{P_1\}\,x{:}A\{P_2\}$, where Δ is a variable context, and P_1,P_2 are unary predicates over m, can be desugared into

$$\mathsf{ST}\,\{\exists\Delta.\,P_1\}\,x{:}A\{\forall\Delta.\,P_1\ \mathsf{i} \supset P_2\ \mathsf{m}\}$$

Similarly, the Hoare type $[\Delta].\,\mathsf{CMD}\,\{I\}\{P_1\}\,x{:}A\{P_2\}$ is desugared into $\mathsf{CMD}\,\{I\}\{\exists\Delta.\,P_1\}\,x{:}A\{\forall\Delta.\,P_1\;\mathsf{i}\supset P_2\;\mathsf{m}\}$. In the rest of the paper, we will use the described convention on ghost variables in order to abbreviate the Hoare types that appear in our examples. However, in the development of the meta theory of HTT, we will assume that postconditions in Hoare types depend on two heap variables: i which denotes the initial heap, and m which denotes the ending heap of the computation.

We call predicates that depend on both i and m *binary predicates*, and use Q and T to range over them. We use X to range over either unary or binary predicates. We will again use the syntax of functional application and write $Q\;H_1\;H_2$ as an abbreviation for $[H_1/\mathsf{i}, H_2/\mathsf{m}]Q$.

We next define several operators on binary predicates that will have a prominent role in the semantics of HTT.

$$
\begin{array}{rcl}
\delta P & = & P \wedge \mathsf{i} = \mathsf{m} \\
\nabla P & = & P \wedge \mathsf{i} = \mathsf{i} \\
\Box P & = & P\;\mathsf{i} \wedge P\;\mathsf{m} \\
X \circ Q & = & \exists h{:}\mathsf{heap}.\,[h/\mathsf{m}]X \wedge Q\;h\;\mathsf{m} \\
Q_1 \ast\!\!\ast\, Q_2 & = & \exists i_1, i_2, m_1, m_2{:}\mathsf{heap}.\;\; \mathsf{splits}(\mathsf{i}, i_1, i_2) \wedge \mathsf{splits}(\mathsf{m}, m_1, m_2) \wedge Q_1\;i_1\;m_1 \wedge Q_2\;i_2\;m_2 \\
P \multimap Q & = & \forall i_0, h{:}\mathsf{heap}.\,\mathsf{splits}(\mathsf{i}, i_0, h) \supset P\;i_0 \supset \exists m_0.\,\mathsf{splits}(\mathsf{m}, m_0, h) \wedge Q\;i_0\;m_0 \\
P_1\,P_2 \multimap\!\!\multimap Q_1\,Q_2 & = & \forall i_0, h{:}\mathsf{heap}.\,\mathsf{splits}(\mathsf{i}, i_0, h) \supset \forall i_1, i_2{:}\mathsf{heap}.\,\mathsf{splits}(i_0, i_1, i_2) \supset \\
 & & \quad P_1\,i_1 \supset P_2\,i_2 \supset \exists m_0, m_1, m_2.\,\mathsf{splits}(m, m_0, h) \wedge \mathsf{splits}(m_0, m_1, m_2) \wedge \\
 & & \quad\quad Q_1\,i_1\,m_1 \wedge Q_2\,i_2\,m_2 \\
M\,?\,Q_1\,Q_2 & = & (M = \mathsf{tt} \supset Q_1) \wedge (M = \mathsf{ff} \supset Q_2)
\end{array}
$$

In English, δP extends the unary predicate P to binary, diagonal one. ∇P is also a binary predicate, albeit one that holds for *any* domain heap (it ignores the variable i). $\Box P$ requires that P holds for both the domain and the range heaps, but unlike δ, does not require that the two heaps are actually equal. $X \circ Q$ is a relational composition. The predicate $Q_1 \ast\!\!\ast\, Q_2$ is the generalization of separating conjunction to the binary case. It holds if both domain and range heaps can be split in two, so that Q_1 relates the first halfs and Q_2 relates the second halfs. $P \multimap Q$ is a binary predicate relating the heap i with m only if m can be obtained by replacing any subheap of i satisfying P with a subheap related by Q. $P_1\,P_2 \multimap\!\!\multimap Q_1\,Q_2$ is the generalization of $P \multimap Q$. It pairwise replaces P_1 according to Q_1 and P_2 according to Q_2 to obtain the heap m starting from i. $M\,?\,Q_1\,Q_2$ is the relational version of a conditional.

Example. The binary relation $(l \mapsto_{\mathsf{nat}} v) \multimap \nabla(l \mapsto_{\mathsf{nat}} v + 1)$ holds between two heaps i and m if and only if m can be obtained from i by replacing *all* parts of i satisfying $l \mapsto_{\mathsf{nat}} v$ (and there can be at most one such part), with a part satisfying $l \mapsto_{\mathsf{nat}} v + 1$. Such a relation therefore directly captures the semantics of an ST computation that increments the contents of l.

4 Type system

The main focus of this section is to describe the techniques used in the definition of the semantics of Hoare types in HTT. The foundations of HTT are in a type theory like the Extended Calculus of Constructions, or Coq, which are very well suited for reasoning about typed functions, relations and sets. In HTT, in addition, we want to support reasoning about effectful computations.

The easiest way to achieve this goal is to *translate* effectful computations into some entity that is already supported by the underlying foundational type theory. In this paper, we have chosen to translate effectful computations into *binary relations on heaps*, so that a computation may be viewed as relating its initial to its ending heap. Choosing relations for the modeling of Hoare types has the additional benefit that we can then also represent partial and non-deterministic computations; that is, computations with no result, or computations with more than one result, respectively.

The translation of computations into relations is performed by the typing rules. Having in mind that there is a strong correspondence between relations and predicates (relations are extensions of predicates),

the reader familiar with Dijkstra's predicate transformers [5] will find the typechecking process completely analogous to computing *strongest postconditions*.

We will have four typing judgments for computations, two for ST-computations and two for CMD-computations. The ST judgments are

$$\Delta; P \vdash E \Rightarrow x{:}A.\, Q$$

and

$$\Delta; P \vdash E \Leftarrow x{:}A.\, Q.$$

The first judgment takes a unary predicate P and a computation E, and generates the binary predicate Q that most tightly captures the semantics of E (i.e., Q is the strongest postcondition). In the process, the rule also verifies that the return result of E has type A, where A is supplied as input to the judgment. The second judgment checks that Q is a postcondition, not necessarily the strongest one for E with respect to P.

The CMD judgments are, similarly,

$$\Delta; I; P \vdash E \Rightarrow x{:}A.\, Q$$

and

$$\Delta; I; P \vdash E \Leftarrow x{:}A.\, Q,$$

except that here P and Q are a pre- and post-condition on the local state of E, while the unary predicate I keeps the invariant on the state that E shares with other processes. By formation, I is required to be precise.

We will make use of further several judgments: (1) $\Delta \vdash K \Rightarrow A$ takes a pure term K and generates its type if it can; (2) $\Delta \vdash M \Leftarrow A$ checks that M has type A. These two judgments implement bidirectional typechecking for the pure fragment. (3) $\Delta \vdash P$ checks that the predicate P is true. It is a completely standard natural deduction for polymorphic first-order logic with equality, except that it also formalizes heaps, via the four axioms listed in Section 3. (4) $\Delta \vdash A \Leftarrow \mathsf{type}$, $\Delta \vdash H \Leftarrow \mathsf{heap}$ and $\Delta \vdash P \Leftarrow \mathsf{prop}$ are type, heap, and predicate formation judgments, and (5) $\Delta \vdash \tau \Leftarrow \mathsf{mono}$ checks that τ is a monomorphic type.

We note one important distinction from the type system in the previous work [20]. In the previous work, we have instrumented the typing judgments so that they always compute the *canonical* (i.e., beta-normal, eta-long) forms of the involved terms. In the current paper, we omit this aspect from the consideration, in order to simplify the presentation. We do note, however, that the equational reasoning is integral to type theory, and an implementation of the system presented here will have to account for it. We are justified in not considering the equational reasoning here, because its details are completely identical in what we have already published in [20]. Indeed, equational reasoning concerns the pure fragment of HTT only, and the pure fragments of the current paper and of [20] are completely identical.

For the sake of illustration, we just list here the beta and eta equations that HTT considers in its equational reasoning. We orient the equations, to emphasize their use as rewriting rules in the normalization algorithm. As customary, beta equality is used as a reduction, and eta equality as an expansion.

For example, the function type $\Pi x{:}A.\, B$ gives rise to the following familiar beta reduction and eta expansion.

$$\begin{array}{rcll} (\lambda x.\, M : \Pi x{:}A.\, B)\ N & \longrightarrow_\beta & [N{:}A/x]M & \\ K & \longrightarrow_\eta & \lambda x.\, K\ x & \text{choosing } x \notin \mathsf{FV}(K) \end{array}$$

Here, of course, the notation $[T/x]M$ denotes a capture-avoiding substitution of the expression T for the variable x in the expression M. In the above equations, the participating terms are decorated with type annotations.

Equations associated with the type $\forall\alpha.\, A$ are also standard.

$$\begin{array}{rcll} (\Lambda\alpha.\, M : \forall\alpha.\, A)\ \tau & \longrightarrow_\beta & [\tau/\alpha]M & \\ K & \longrightarrow_\eta & \Lambda\alpha.\, K\ \alpha & \text{choosing } \alpha \notin \mathsf{FTV}(K) \end{array}$$

where $\mathsf{FTV}(K)$ denotes the free monotype variables of K.

In the case of the unit type, we do not have any beta reduction (as there are no elimination constructors associated with this type), but only one eta expansion.

$$K \quad \longrightarrow_\eta \quad ()$$

The equations for the Hoare type require an auxiliary operation of *monadic substitution*, $\langle E/x{:}A\rangle F$, which sequentially composes the computations E and F. The result is a computation which should, intuitively, evaluate as E *followed by* F, where the free variable $x{:}A$ in F is bound to the value of E. Monadic substitution is defined by induction on the structure of E, as follows.

$$\begin{array}{rcll}
\langle \mathsf{return}\ M/x{:}A\rangle F & = & [M{:}A/x]F & \\
\langle y \Leftarrow K; E/x{:}A\rangle F & = & y \Leftarrow K; \langle E/x{:}A\rangle F & \text{choosing } y \notin \mathsf{FV}(F) \\
\langle y \Leftarrow !_\tau\, M; E/x{:}A\rangle F & = & y \Leftarrow !_\tau\, M; \langle E/x{:}A\rangle F & \text{choosing } y \notin \mathsf{FV}(F) \\
\langle M :=_\tau N; E/x{:}A\rangle F & = & M :=_\tau N; \langle E/x{:}A\rangle F & \\
\langle y \Leftarrow \mathsf{alloc}_\tau\, M; E/x{:}A\rangle F & = & y \Leftarrow \mathsf{alloc}_\tau\, M; \langle E/x{:}A\rangle F & \text{choosing } y \notin \mathsf{FV}(F) \\
\langle \mathsf{dealloc}\, M; E/x{:}A\rangle F & = & \mathsf{dealloc}\, M; \langle E/x{:}A\rangle F & \\
\langle y \Leftarrow \mathsf{atomic}_{B,P,Q}\, E_1; E/x{:}A\rangle F & = & y \Leftarrow \mathsf{atomic}_{B,P,Q}\, E_1; \langle E/x{:}A\rangle F & \text{choosing } y \notin \mathsf{FV}(F) \\
\langle (y_1{:}A_1 \Leftarrow E_1{:}P_1 \parallel & = & (y_1{:}A_1 \Leftarrow E_1{:}P_1 \parallel & \text{choosing } y_1, y_2 \notin \mathsf{FV}(F) \\
\quad y_2{:}A_2 \Leftarrow E_2{:}P_2); E/x{:}A\rangle F & & \quad y_2{:}A_2 \Leftarrow E_2{:}P_2); \langle E/x{:}A\rangle F & \\
\langle \mathsf{publish}\ J; E/x{:}A\rangle F & = & \mathsf{publish}\ J; \langle E/x{:}A\rangle F & \\
\langle y \Leftarrow \mathsf{fix}_A\, M; E/x{:}A\rangle F & = & y \Leftarrow \mathsf{fix}_A\, M; \langle E/x{:}A\rangle F & \text{choosing } y \notin \mathsf{FV}(F) \\
\langle y \Leftarrow \mathsf{if}_A\ M\ \mathsf{then}\ E_1\ \mathsf{else}\ E_2; E/x{:}A\rangle F & = & y \Leftarrow \mathsf{if}_A\ M\ \mathsf{then}\ E_1\ \mathsf{else}\ E_2; \langle E/x{:}A\rangle F & \text{choosing } y \notin \mathsf{FV}(F)
\end{array}$$

Now the equations associated with Hoare types can be defined as follows.

$$\begin{array}{rcl}
x \leftarrow (\mathsf{stdo}\ E : \mathsf{ST}\,\{P\}\,x{:}A\{Q\}); F & \longrightarrow_\beta & \langle E/x{:}A\rangle F \\
K & \longrightarrow_\eta & \mathsf{stdo}\ (x \leftarrow K; \mathsf{return}\ x)
\end{array}$$

$$\begin{array}{rcl}
x \leftarrow (\mathsf{cmdo}\ E : \mathsf{CMD}\,\{I\}\{P\}\,x{:}A\{Q\}); F & \longrightarrow_\beta & \langle E/x{:}A\rangle F \\
K & \longrightarrow_\eta & \mathsf{cmdo}\ (x \leftarrow K; \mathsf{return}\ x)
\end{array}$$

The monadic substitution is modeled directly after the operation introduced by Pfenning and Davies [26], who also show that the beta and eta equations above are equivalent to the unit and associativity laws for a generic monad [17].

The monadic equations associated with the ST-monad are completely analogous to the ones for the CMD-monad, and possess the same properties, which for the case of the ST monads, we have considered in [20].

There are several more reductions that factor into definitional equality. For example, $\bar{m} \oplus \bar{n}$ is not canonical, as it can be simplified into $\overline{m \oplus n}$. These reductions are required in order for the normalization to agree with the evaluation of nat expressions, so that the actual numerals are the only closed canonical forms of type nat.

4.1 Type and mono formation rules.

$$\frac{}{\Delta, \alpha, \Delta_1 \vdash \alpha \Leftarrow \mathsf{type}} \qquad \frac{}{\Delta \vdash \mathsf{bool} \Leftarrow \mathsf{type}} \qquad \frac{}{\Delta \vdash \mathsf{nat} \Leftarrow \mathsf{type}} \qquad \frac{}{\Delta \vdash 1 \Leftarrow \mathsf{type}}$$

$$\frac{\Delta \vdash A \Leftarrow \mathsf{type} \qquad \Delta, x{:}A \vdash B \Leftarrow \mathsf{type}}{\Delta \vdash \Pi x{:}A.\, B \Leftarrow \mathsf{type}} \qquad \frac{\Delta, \alpha \vdash A \Leftarrow \mathsf{type}}{\Delta \vdash \forall \alpha.\, A \Leftarrow \mathsf{type}} \qquad \frac{\begin{array}{c}\Delta, \mathsf{m{:}heap} \vdash P \Leftarrow \mathsf{prop} \qquad \Delta \vdash A \Leftarrow \mathsf{type} \\ \Delta, x{:}A, \mathsf{i{:}heap}, \mathsf{m{:}heap} \vdash Q \Leftarrow \mathsf{prop}\end{array}}{\Delta \vdash \mathsf{ST}\,\{P\}\,x{:}A\{Q\} \Leftarrow \mathsf{type}}$$

$$\frac{\begin{array}{c}\Delta, \mathsf{m{:}heap} \vdash I \Leftarrow \mathsf{prop} \qquad \Delta, \mathsf{m{:}heap} \vdash \mathsf{precise}\ I \qquad \Delta, \mathsf{m{:}heap} \vdash P \Leftarrow \mathsf{prop} \\ \Delta \vdash A \Leftarrow \mathsf{type} \qquad \Delta, x{:}A, \mathsf{i{:}heap}, \mathsf{m{:}heap} \vdash Q \Leftarrow \mathsf{prop}\end{array}}{\Delta \vdash \mathsf{CMD}\,\{I\}\{P\}\,x{:}A\{Q\} \Leftarrow \mathsf{type}}$$

We omit the rules for the mono judgment, but they are obvious. That is, they just repeat the rules for the type formation, except that the rule for universal quantification is omitted.

4.2 Heap formation rules.

$$\frac{}{\Delta, h{:}\mathsf{heap}, \Delta_1 \vdash h \Leftarrow \mathsf{heap}} \qquad \frac{}{\Delta \vdash \mathsf{empty} \Leftarrow \mathsf{heap}}$$

$$\frac{\Delta \vdash \tau \Leftarrow \mathsf{mono} \qquad \Delta \vdash H \Leftarrow \mathsf{heap} \qquad \Delta \vdash M \Leftarrow \mathsf{nat} \qquad \Delta \vdash N \Leftarrow \tau}{\Delta \vdash \mathsf{upd}_\tau(H, M, N) \Leftarrow \mathsf{heap}}$$

4.3 Prop formation rules.

$$\frac{\Delta \vdash A \Leftarrow \mathsf{type} \qquad \Delta \vdash M \Leftarrow A \qquad \Delta \vdash N \Leftarrow A}{\Delta \vdash \mathsf{id}_A(M, N) \Leftarrow \mathsf{prop}}$$

$$\frac{\Delta \vdash \tau \Leftarrow \mathsf{mono} \qquad \Delta \vdash H \Leftarrow \mathsf{heap} \qquad \Delta \vdash M \Leftarrow \mathsf{nat} \qquad \Delta \vdash N \Leftarrow \tau}{\Delta \vdash \mathsf{seleq}_\tau(H, M, N) \Leftarrow \mathsf{prop}} \qquad \frac{}{\Delta \vdash \top \Leftarrow \mathsf{prop}}$$

$$\frac{}{\Delta \vdash \bot \Leftarrow \mathsf{prop}} \qquad \frac{\Delta \vdash P \Leftarrow \mathsf{prop} \qquad \Delta \vdash Q \Leftarrow \mathsf{prop}}{\Delta \vdash P \wedge Q \Leftarrow \mathsf{prop}} \qquad \frac{\Delta \vdash P \Leftarrow \mathsf{prop} \qquad \Delta \vdash Q \Leftarrow \mathsf{prop}}{\Delta \vdash P \vee Q \Leftarrow \mathsf{prop}}$$

$$\frac{\Delta \vdash P \Leftarrow \mathsf{prop} \qquad \Delta \vdash Q \Leftarrow \mathsf{prop}}{\Delta \vdash P \supset Q \Leftarrow \mathsf{prop}} \qquad \frac{\Delta \vdash P \Leftarrow \mathsf{prop}}{\Delta \vdash \neg P \Leftarrow \mathsf{prop}} \qquad \frac{\Delta \vdash A \Leftarrow \mathsf{type} \qquad \Delta, x{:}A \vdash P \Leftarrow \mathsf{prop}}{\Delta \vdash \forall x{:}A.\, P \Leftarrow \mathsf{prop}}$$

$$\frac{\Delta \vdash A \Leftarrow \mathsf{type} \qquad \Delta, x{:}A \vdash P \Leftarrow \mathsf{prop}}{\Delta \vdash \exists x{:}A.\, P \Leftarrow \mathsf{prop}} \qquad \frac{\Delta, h{:}\mathsf{heap} \vdash P \Leftarrow \mathsf{prop}}{\Delta \vdash \forall h{:}\mathsf{heap}.\, P \Leftarrow \mathsf{prop}} \qquad \frac{\Delta, h{:}\mathsf{heap} \vdash P \Leftarrow \mathsf{prop}}{\Delta \vdash \exists h{:}\mathsf{heap}.\, P \Leftarrow \mathsf{prop}}$$

$$\frac{\Delta, \alpha \vdash P \Leftarrow \mathsf{prop}}{\Delta \vdash \forall \alpha.\, P \Leftarrow \mathsf{prop}} \qquad \frac{\Delta, \alpha \vdash P \Leftarrow \mathsf{prop}}{\Delta \vdash \exists \alpha.\, P \Leftarrow \mathsf{prop}}$$

4.4 Pure terms.

Because monads hygienically isolate the effectful computations from the pure ones, the pure fragment may be formulated in a standard way, drawing on the formalism of ECC.

$$\frac{}{\Delta, x{:}A, \Delta_1 \vdash x \Rightarrow A} \qquad \frac{\Delta, x{:}A \vdash M \Leftarrow B}{\Delta \vdash \lambda x.\, M \Leftarrow \Pi x{:}A.\, B} \qquad \frac{\Delta \vdash K \Rightarrow \Pi x{:}A.\, B \qquad \Delta \vdash M \Leftarrow A}{\Delta \vdash K\ M \Rightarrow [M{:}A/x]B}$$

$$\frac{\Delta, \alpha \vdash M \Leftarrow A}{\Delta \vdash \Lambda\alpha.\, M \Leftarrow \forall \alpha.\, A} \qquad \frac{\Delta \vdash K \Rightarrow \forall \alpha.\, A \qquad \Delta \vdash \tau \Leftarrow \mathsf{mono}}{\Delta \vdash K\ \tau \Rightarrow [\tau/\alpha]A} \qquad \frac{\Delta \vdash K \Rightarrow B \qquad A = B}{\Delta \vdash K \Leftarrow A} \qquad \frac{\Delta \vdash M \Leftarrow A}{\Delta \vdash M{:}A \Rightarrow A}$$

In HTT, we adopt a formulation with bidirectional typechecking, inspired by the work on Concurrent LF [30], where elimination terms synthesize their types, and introduction terms must be supplied a type to check against. We can always check an elimination term K by first synthesizing its type and confirming that it is equal with the supplied type. We can always switch the direction from checking to synthesis by using the constructor $M{:}A$ to supply the type A to be synthesized.

4.5 Assertion Logic

The assertion logic implements the standard natural deduction for a first-order predicate logic. The operators for truth, falsehood, conjunction, discjunction, implication, and quantifiers are formulated in the usual way.

For the sake of concreteness, we only present here the rules that axiomatize equality, function extensionality, heaps and booleans. Natural numbers are axiomatized by the usual Peano axioms. We start with equality:

$$\frac{\Delta \vdash M \Leftarrow A}{\Delta \vdash \mathsf{id}_A(M, M)} \qquad \frac{\Delta \vdash \mathsf{id}_A(M, N) \qquad \Delta, x{:}A \vdash P \Leftarrow \mathsf{prop} \qquad \Delta \vdash [M{:}A/x]P}{\Delta \vdash [N{:}A/x]P}$$

It is well-known [10], that the equality rules above do not admit extensional equality of functions. The terms M and N must depend only on the variables in Δ, while extensional equality of functions require extending the context with an additional variable. We require a separate rule for function extensionality, and similarly, a separate rule for equality of type abstractions.

$$\frac{\Delta, x{:}A \vdash \mathsf{id}_B(M, N)}{\Delta \vdash \mathsf{id}_{\Pi x{:}A.\, B}(\lambda x.\, M, \lambda x.\, N)} \qquad \frac{\Delta, \alpha \vdash \mathsf{id}_B(M, N)}{\Delta \vdash \mathsf{id}_{\forall \alpha.\, B}(\Lambda \alpha.\, M, \Lambda \alpha.\, N)}$$

In the above rules, it is assumed that the bound variables x and α do not appear free in the involved contexts.

Heaps are axiomatized using the rules that we paraphrased and commented on in Section 3. Here we put them in the proper formal setting, as inference rules of the judgment $\Delta \vdash P$.

$$\frac{}{\Delta \vdash \mathsf{seleq}_\tau(\mathsf{empty}, M, N)} \qquad \frac{}{\Delta \vdash \mathsf{seleq}_\tau(\mathsf{upd}_\tau(H, M, N), M, N)}$$

$$\frac{\Delta \vdash \neg \mathsf{id}_{\mathsf{nat}}(M_1, M_2) \qquad \Delta \vdash \mathsf{seleq}_\tau(\mathsf{upd}_\sigma(H, M_1, N_1), M_2, N_2)}{\Delta \vdash \mathsf{seleq}_\tau(H, M_2, N_2)}$$

$$\frac{\Delta \vdash \mathsf{seleq}_\tau(H, M, N_1) \qquad \Delta \vdash \mathsf{seleq}_\tau(H, M, N_2)}{\Delta \vdash \mathsf{id}_\tau(N_1, N_2)}$$

Rules for booleans need to state that tt and ff are different. They also need to state an extensionality principle, by which a boolean can only be equal to tt or ff.

$$\frac{}{\Delta \vdash \neg \mathsf{id}_{\mathsf{bool}}(\mathsf{tt}, \mathsf{ff})} \qquad \frac{\Delta, x{:}\mathsf{bool} \vdash P \Leftarrow \mathsf{prop} \qquad \Delta \vdash [\mathsf{tt}{:}\mathsf{bool}/x]P \qquad \Delta \vdash [\mathsf{ff}{:}\mathsf{bool}/x]P \qquad \Delta \vdash M \Leftarrow \mathsf{bool}}{\Delta \vdash [M{:}\mathsf{bool}/x]P}$$

4.6 Typechecking ST-computations

We start with a structural rule which relates the synthesis and checking of postconditions: if Q' is a strongest postcondition, and from knowing Q' we can derive Q, then Q is a postcondition.

$$\frac{\Delta; P \vdash E \Rightarrow x{:}A.\, Q' \qquad \Delta, x{:}A, \mathsf{i}, \mathsf{m}{:}\mathsf{heap}, \delta P \circ Q' \vdash Q}{\Delta; P \vdash E \Leftarrow x{:}A.\, Q}$$

Rather than simply taking Q' as a hypothesis when trying to derive Q, the rule takes $\delta P \circ Q'$. Unrolling the definitions of $\circ$ and δ, this basically injects the knowledge that the initial heap of Q also satisfies P, which should be expected as P is a precondition that the checking starts with. This rule essentially implements the law of consequence well-known in Hoare Logic.

The typing rule for monadic unit in a sense corresponds to a rule for assignment to the variable x found in the classical formulations of Hoare Logic:

$$\frac{\Delta \vdash M \Leftarrow A}{\Delta; P \vdash \mathsf{return}\ M \Rightarrow x{:}A.\, \delta(x = M)}$$

The postcondition $\delta(x = M)$ simply states that after executing $\mathsf{return}\ M$ the return value $x = M$ and the initial and the ending heap are equal since no state has been modified.

$$\frac{\Delta \vdash \tau \Leftarrow \mathsf{mono} \qquad \Delta \vdash M \Leftarrow \mathsf{nat} \qquad \Delta, \mathsf{m}{:}\mathsf{heap}, P \vdash M \hookrightarrow_\tau - \qquad \Delta, x{:}\tau; P \circ \delta(M \hookrightarrow_\tau x) \vdash E \Rightarrow y{:}B.\, Q}{\Delta; P \vdash x \Leftarrow !_\tau M; E \Rightarrow y{:}B.\, (\exists x{:}\tau.\, \delta(M \hookrightarrow_\tau x) \circ Q)}$$

The rule for memory read must check that τ is a well-formed monomorphic type, as only values of monomorphic types can be stored into heaps. Further, the location M must be a natural number, and M must point to a value of type τ. The later is ensured by the entailment $P \vdash M \hookrightarrow_\tau -$, which may be seen as a *verification condition* that needs to be discharged in order to guarantee the correctness of the program. The continuation E is then checked in a context extended with variable $x{:}\tau$, and the precondition for checking E must appropriately reflect the knowledge that x binds the value read from M. This is achieved by composing P with $\delta(M \hookrightarrow_\tau x)$. Alternatively, we could have used the equivalent $P \wedge (M \hookrightarrow_\tau x)$, which is the standard postcondition for memory lookup, but we choose the current formulation in order to emphasize the compositional nature of typechecking. For example, after the relation Q corresponding to E is obtained, we need to lift it to include the semantics of the lookup, before we return it as a postcondition generated for the original computation. We do so by composing $\delta(M \hookrightarrow_\tau x) \circ Q$. One can now see the important intuition that, in general, the strongest postcondition generated for some computation E always has a form of an ordered sequence of compositions of smaller relations, each of which precisely captures the primitive effectful commands of E, in the order in which they appear in E. This substantiates our claim that typechecking simply translates the E into a relation (equivalently, predicate). In fact, the translation is almost literal, as the structure of the obtained predicate completely mirrors E.

Memory writes follow a similar strategy.

$$\frac{\Delta \vdash M \Leftarrow \mathsf{nat} \qquad \Delta \vdash \tau \Leftarrow \mathsf{mono} \qquad \Delta \vdash N \Leftarrow \tau \qquad \Delta, \mathsf{m{:}heap}, P \vdash M \hookrightarrow - \qquad \Delta; P \circ (M \mapsto - \multimap \nabla(M \mapsto_\tau N)) \vdash E \Rightarrow x{:}A.\, Q}{\Delta; P \vdash M :=_\tau N; E \Rightarrow x{:}A.\, (M \mapsto - \multimap \nabla(M \mapsto_\tau N)) \circ Q}$$

To write into the location M, we first ensure that it is allocated (verification condition $P \vdash M \hookrightarrow -$). Then the continuation E is checked with respect to a predicate $P \circ (M \mapsto - \multimap \nabla(M \mapsto_\tau N))$. Intuitively, following the definition of the connective $\multimap$, this predicate "replaces" a portion of the heap satisfying $M \mapsto -$ by a heap satisfying $M \mapsto_\tau N$, while preserving the rest of the structure described by P. Thus, the predicate correctly models the semantics of memory lookup.

The idea behind the typechecking of stateful commands should now be obvious, so we simply display the rules for allocation and deallocation without further comment.

$$\frac{\Delta \vdash \tau \Leftarrow \mathsf{mono} \qquad \Delta \vdash M \Leftarrow \tau \qquad \Delta, x{:}\tau; P \circ (\mathsf{emp} \multimap \nabla(x \mapsto_\tau M)) \vdash E \Rightarrow y{:}B.\, Q}{\Delta; P \vdash x \Leftarrow \mathsf{alloc}_\tau\, M; E \Rightarrow y{:}B.\, (\exists x{:}\tau.\, (\mathsf{emp} \multimap \nabla(x \mapsto_\tau M)) \circ Q)}$$

$$\frac{\Delta \vdash M \Leftarrow \mathsf{nat} \qquad \Delta, \mathsf{m{:}heap}, P \vdash M \hookrightarrow - \qquad \Delta; P \circ (M \mapsto - \multimap \nabla \mathsf{emp}) \vdash E \Rightarrow x{:}A.\, Q}{\Delta; P \vdash \mathsf{dealloc}\, M; E \Rightarrow x{:}A.\, (M \mapsto - \multimap \nabla \mathsf{emp}) \circ Q}$$

4.7 Typechecking CMD-computations

The CMD-judgments have similar rules for consequence and unit as the one presented in the ST case. We omit these here, and focus instead on the primitives for concurrency.

$$\frac{\Delta \vdash A_i \Leftarrow \mathsf{type} \qquad \Delta, \mathsf{m{:}heap} \vdash P_i \Leftarrow \mathsf{prop} \qquad \Delta; \mathsf{m{:}heap}, P \vdash P_1 * P_2 * \top \qquad \Delta; I; P_1 \vdash E_1 \Rightarrow x_1{:}A_1.\, Q_1 \qquad \Delta; I; P_2 \vdash E_2 \Rightarrow x_2{:}A_2.\, Q_2 \qquad \Delta, x_1{:}A_1, x_2{:}A_2; I; P \circ (P_1\, P_2 \multimap\!\!\multimap Q_1\, Q_2) \vdash E \Rightarrow x{:}A.\, Q}{\Delta; I; P \vdash (x_1{:}A_1 \Leftarrow E_1{:}P_1 \parallel x_2{:}A_2 \Leftarrow E_2{:}P_2); E \Rightarrow x{:}A.\, (\exists x_1{:}A_1, x_2{:}A_2.\, (P_1\, P_2 \multimap\!\!\multimap Q_1\, Q_2) \circ Q)}$$

The command $\parallel$ for fork-join parallelism checks the processes E_1 and E_2 with respect to the ascribed preconditions on local state P_1 and P_2 to obtain strongest postconditions Q_1 and Q_2. A verification condition is issued to check that P_1 and P_2 indeed split the local state of the parent process into disjoint sections (the entailment $P \vdash P_1 * P_2 * \top$). Then the common continuation E is checked with respect to a new description of the state $P \circ (P_1\, P_2 \multimap\!\!\multimap Q_1\, Q_2)$, which captures the semantics that the local heap described by P is changed so that the P_i fragment are independently updated according to Q_i. Thus, the predicate correctly captures

the semantics of concurrent execution of E_1 and E_2. Since E_1 and E_2 are checked using the same shared invariant I, they can modify the shared state, but only in ways which preserve the truth value of I.

$$\frac{\begin{array}{c}\Delta \vdash A_1 \Leftarrow \mathsf{type} \qquad \Delta, \mathsf{m{:}heap} \vdash P_1 \Leftarrow \mathsf{prop} \qquad \Delta, \mathsf{m{:}heap}, \mathsf{i{:}heap} \vdash Q_1 \Leftarrow \mathsf{prop} \\ \Delta, \mathsf{m{:}heap}, P \vdash P_1 * \top \qquad \Delta; I * P_1 \vdash E_1 \Leftarrow x_1{:}A_1.\, \Box I \mathbin{*\!\!*} Q_1 \qquad \Delta, x_1{:}A_1; I; P \circ (P_1 \multimap Q_1) \vdash E \Rightarrow x{:}A.\, Q\end{array}}{\Delta; I; P \vdash x_1 \Leftarrow \mathsf{atomic}_{A_1,P_1,Q_1} E_1; E \Rightarrow x{:}A.\, (\exists x_1{:}A_1.\, (P_1 \multimap Q_1) \circ Q)}$$

Semantically, atomic first acquires exclusive access to the shared state, executes E_1 using both the shared and the designated chunk of the local state. Thus, E_1 must be checked against a precondition $I * P_1$, where I is the descriptor of the shared state, and P_1 is the descriptor of the designated chunk of the local state. It is ensured that P_1 describes local state by emitting a verification condition $P \vdash P_1 * \top$. We emphasize that E_1 is an ST-computation, and thus it makes no semantic distinction between local and shared state. Upon exit, E_1 releases what used to be shared state, so it must make sure that its invariant is preserved. Thus, E_1 is checked against a postcondition $\Box I \mathbin{*\!\!*} Q$, which requires that E_1 changes the I portion of the its initial heap in such a way that the changed subheap satisfies I again. This portion is what will be released as an updated version of the shared state. The rest of the heap is what used to be the local state of the parent process, and is changed according to some pre-specified Q. The continuation E is then simply checked against a local heap that updates a P_1 part of P according to the binary relation Q. The update is expressed using the relation $P \circ (P_1 \multimap Q_1)$.

In order for the semantics to make sense, we must make sure that there is only one portion in the combined shared/local heap that satisfies I, else E_1 may not know how much space must be returned to shared status. That is why I is required to be precise, enforcing that it always determines a unique subheap. Precision is a standard requirement on invariants of shared resources in Separation Logic [3, 22, 29].

$$\frac{\begin{array}{c}\Delta, \mathsf{m{:}heap} \vdash J \Leftarrow \mathsf{prop} \\ \Delta, \mathsf{m{:}heap} \vdash \mathsf{precise}\ J \qquad \Delta, \mathsf{m{:}heap}, P \vdash J * \top \qquad \Delta; I * J; P \circ (J \multimap \nabla\mathsf{emp}) \vdash E \Rightarrow x{:}A.\, Q\end{array}}{\Delta; I; P \vdash \mathsf{publish}\ J; E \Rightarrow x{:}A.\, (J \multimap \nabla\mathsf{emp}) \circ Q}$$

Publish takes a predicate J and promotes the chunk of the local heap satisfying J into shared status. Thus, J must hold of a unique part of the local heap. Existence of such a part is ensured by the verification condition $P \vdash J * \top$, and uniqueness is ensured by the requirement that J is precise. The published state is shared throughout the scope of the continuation E, which must be checked against an extended invariant on the shared state $(I * J)$ and a description of a shrunken local state $P \circ (J \multimap \nabla\mathsf{emp})$. The later predicate simply states that the unique J part of P is replaced by an empty heap, thus subtracting J from P.

4.8 Typechecking generic computational primitives

The typing rule for conditional is unsurprising; it obtains the postconditions for the branches, and then checks the continuation with what amounts to a disjunction of these postconditions. We present only the CMD rule, as the ST rule is analogous.

$$\frac{\begin{array}{c}\Delta \vdash A \Leftarrow \mathsf{type} \qquad \Delta \vdash M \Leftarrow \mathsf{bool} \qquad \Delta; I; P \circ \delta(M = \mathsf{tt}) \vdash E_1 \Rightarrow x{:}A.\, Q_1 \\ \Delta; I; P \circ \delta(M = \mathsf{ff}) \vdash E_2 \Rightarrow x{:}A.\, Q_2 \qquad \Delta, x{:}A; I; P \circ (M\,?\,Q_1\,Q_2) \vdash E \Rightarrow y{:}B.\, Q\end{array}}{\Delta; I; P \vdash x \Leftarrow \mathsf{if}_A\ M\ \mathsf{then}\ E_1\ \mathsf{else}\ E_2; E \Rightarrow y{:}B.\, (\exists x{:}A.\, (M\,?\,Q_1\,Q_2) \circ Q)}$$

The fragment of HTT described so far may easily be presented in a more customary form with Hoare triples for partial correctness, because the constructs have been essentially first-order. We now describe the two effectful constructs which are higher-order in an essential way. They have well-known first-order analogues, but these are significantly less expressive.

The first is monadic bind, whose first-order analogue is the Hoare rule for sequential composition.

$$\frac{\begin{array}{c}\Delta \vdash K \Rightarrow \mathsf{CMD}\,\{I_1\}\{P_1\}\,x_1{:}A_1\{Q_1\} \\ \Delta, \mathsf{m{:}heap}, I \vdash I_1 * (I_1 \mathbin{-\!\!*} I) \qquad \Delta, \mathsf{m{:}heap}, P \vdash P_1 * \top \qquad \Delta, x_1{:}A_1; I; P \circ (P_1 \multimap Q_1) \vdash E \Rightarrow x{:}A.\, Q\end{array}}{\Delta; I; P \vdash x_1 \Leftarrow K; E \Rightarrow x{:}A.\, (\exists x_1{:}A_1.\, (P_1 \multimap Q_1) \circ Q)}$$

The difference is that monadic bind allows the first composed computation to be obtained by *evaluating* K, whereas in Hoare Logic, the composed processes must be supplied explicitly. HTT, as well as other monadic calculi, treats computations as first-class values which can be supplied as function arguments, obtained as function results, and abstracted. These features are the essential ingredients for our lifting of Hoare Logic to higher-order.

Returning to the specifics of bind, we notice that the code encapsulated by K is executable in a local heap satisfying P (ensured by the verification condition $P \vdash P_1 * \top$), and a shared heap satisfying I (ensured by the verification condition $I \vdash I_1 * (I_1 \mathrel{-\!\!*} I)$). The later condition implies that I_1 is a *restriction* of I to the subheap determined by I_1, thus ensuring that K, by preserving the invariant I_1, also preserves the larger invariant I. Finally, according to the Hoare type of K, its execution changes the P_1 chunk of the local heap as described by the postcondition Q_1. Thus, the continuation E is appropriately checked against a precondition $P \circ (P_1 \multimap Q_1)$.

The second higher-order connective is the fixpoint. Due to the presence of higher-order functions, HTT can replace the looping constructs by a general fixpoint combinator over a Hoare type. The fixpoint computation is supposed to be immediately executed, so the typing rule combines the usual typing of fixpoint combinators with monadic bind.

$$\frac{\begin{array}{c} B = \mathsf{CMD}\,\{I_1\}\{P_1\}\,x_1{:}A_1\{Q_1\} \qquad \Delta \vdash B \Leftarrow \mathsf{type} \qquad \Delta, \mathsf{m{:}heap}, I \vdash I_1 * (I_1 \mathrel{-\!\!*} I) \\ \Delta, \mathsf{m{:}heap}, P \vdash P_1 * \top \qquad \Delta \vdash f \Leftarrow B \to B \qquad \Delta, x_1{:}A_1; I; P \circ (P_1 \multimap Q_1) \vdash E \Rightarrow x{:}A.\, Q \end{array}}{\Delta; I; P \vdash x_1 \Leftarrow \mathsf{fix}_B f; E \Rightarrow x{:}A.\, (\exists x_1{:}A_1.\, (P_1 \multimap Q_1) \circ Q)}$$

The ST-judgment contains similar rules for monadic bind and fixpoint. They are strictly simpler than the above, as they do not have to account for the the shared state invariant I.

Finally, we integrate the monadic judgments into the pure fragment, using the two do coercions.

$$\frac{\Delta; P \vdash E \Leftarrow x{:}A.\, Q}{\Delta \vdash \mathsf{stdo}\ E \Leftarrow \mathsf{ST}\,\{P\}\,x{:}A\{Q\}}$$

$$\frac{\Delta; I; P \vdash E \Leftarrow x{:}A.\, Q}{\Delta \vdash \mathsf{cmdo}\ E \Leftarrow \mathsf{CMD}\,\{I\}\{P\}\,x{:}A\{Q\}}$$

4.9 Annotated example

The code in Figure 1 presents the typing derivation of the function $\mathsf{waitThen}$ from Section 2.1. We check the function against the type:

$$\forall\alpha.\, \Pi l{:}\mathsf{loc}.\, \Pi n{:}\mathsf{nat}.\, \Pi r{:}\mathsf{loc}.\, \mathsf{ST}\{(l \mapsto_{\mathsf{nat}} n) * J * P\}x{:}\alpha\{(l \mapsto_{\mathsf{nat}} -) * J * Q(x)\} \to A$$

where

$$A = \mathsf{CMD}\{(l \mapsto_{\mathsf{nat}} -) * J\}\{P * (r \mapsto_{\mathsf{nat}} 0)\}t{:}1\{\exists x{:}\alpha.\, Q(x) * (r \mapsto_{\alpha} x)\}.$$

The code is annotated with predicates to illustrate the properties of the state at the various program points. We use I for predicates about the shared state, and P for predicates about local state. The typechecker will require that several verification conditions be proved, before the program is deemed correct. These are: (1) $P_2 \vdash P * (r \mapsto_{\mathsf{nat}} 0) * \top$, to enter atomic; (2) $P_3 \vdash l \hookrightarrow_{\mathsf{nat}} -$, to read from l; (3) $P_5 \vdash (l \mapsto_{\mathsf{nat}} n) * J * P * \top$, to execute st; (4) $P_6 \vdash r \hookrightarrow -$, to write into r; (5) $P_8 \vdash \Box I_2 \mathbin{**} (t_1\,?\,(P * (r \mapsto_{\mathsf{nat}} 0))\,(\exists y{:}\alpha.\, Q(y) * (r \mapsto_{\alpha} y)))$, to prove that the postcondition supplied as an index to atomic is satisfied, and thus atomic exits correctly; (6) $I_{10} \vdash ((l \mapsto_{\mathsf{nat}} -) * J) * (((l \mapsto_{\mathsf{nat}} -) * J) \mathrel{-\!\!*} I_{10})$ and $P_{10} \vdash P * (r \mapsto_{\mathsf{nat}} 0) * \top$, to execute the recursive call to c, and (7) $P_{12} \vdash \exists x{:}\alpha.\, Q(x) * (r \mapsto_{\alpha} x)$, to prove that the postcondition of $\mathsf{waitThen}$ holds. All of these conditions are fairly easy to discharge.

```
waitThen = Λα. λl. λn. λst. λr.
  cmdo(-- I1:{(l ↦nat −) ∗ J}
       -- P1:{P ∗ r ↦nat 0}
       t ⇐ fixA(λc. cmdo
         (-- I2:{(l ↦nat −) ∗ J}
          -- P2:{P ∗ r ↦nat 0}
          ok ⇐ atomic bool, P ∗ (r ↦nat 0),
                        t1. (t1 ? (P ∗ r ↦nat 0)
                                  (∃y:α. Q(y) ∗ r ↦α y))
                 (-- P3:{l ↦nat − ∗ J ∗ P2}
                  x ⇐ !nat l;
                  -- P4:{l ↦nat x ∗ J ∗ P2}
                  t1 ⇐  if (x = n) then
                          -- P5:{l ↦nat n ∗ J ∗ P2}
                          y ⇐ st;
                          -- P6:{l ↦nat − ∗ J ∗ Q(y) ∗ r ↦nat 0}
                          r :=α y;
                          -- P7:{l ↦nat − ∗ J ∗ Q(y) ∗ r ↦α y}
                          return ff
                          else return tt;
                  -- P8:{(t1 = ff ⊃ ∃y:α. P7) ∧ (t1 = tt ⊃ P4}
                  return t1);
          -- I9:{(l ↦nat −) ∗ J}
          -- P9:{ok ? (P ∗ r ↦nat 0)
                      (∃y:α. Q(y) ∗ r ↦α y)}
          if ok then
               -- I10:{(l ↦nat −) ∗ J}
               -- P10:{P ∗ r ↦nat 0}
               t3 ⇐ c;
               -- I11:{(l ↦nat −) ∗ J}
               -- P11:{∃y:α. Q(y) ∗ r ↦α y}
               return t3
          else return ())));
       -- I12:{(l ↦nat −) ∗ J}
       -- P12:{∃y:α. Q(y) ∗ r ↦α y}
       return t)
```

Figure 1: Annotated example.

5 Properties

5.1 Equational reasoning and logical soundness

Like every other type theory, HTT has to define a judgment $A = B$ for checking if two types A and B are equal. This judgment was used, for example, in the typing rule of the pure fragment where the bidirectional typechecker switches from synthesis to checking mode. Because types are dependent, and thus may contain terms, the judgment is non-trivial, as it has to account for term equality as well. In addition, it is a desirable property that the equality judgment be decidable.

Thus, the equality judgment of HTT, as of other (intensional) type theories, like ECC or Coq, is quite restricted and includes only equations that lead to decidable theories. In Coq, for example, the equality judgment only admits beta reduction. Other equations of interest may, of course, be added as axioms. Properties that rely on such axioms cannot be proved automatically by the typechecker, but must be discharged by the user via explicit proofs.

The development of the monadic fragment of HTT does not depend on the organization and the equations of the pure fragment. So for example, if we chose Coq as an environment type theory of HTT, we could freely use beta reduction in the equational reasoning.

In the previous work [20], we have allowed equational reasoning that relies on beta and eta equalities for Π-types, and the generic monadic laws (unit and associativity [17]) for Hoare types. We have shown that such an equational theory is decidable, using the technically involved, but conceptually quite simple and elegant idea of *hereditary substitutions* [30]. The current paper uses literally the same pure fragment, so the same proof of decidability applies.

Theorem 1 (Relative decidability of typechecking) *Given an oracle for deciding the verification conditions (that is, deciding the judgment $\Delta \vdash P$, where P is a proposition), all the typing judgments of HTT are decidable.*

In the actual implementation of HTT, the oracle from the above theorem can be replaced by explicit proofs, to be obtained by some form of automatic or interactive theorem proving. The later has been the approach that we adopted in the implementation of HTT in Coq, where a modicum of automation can be recovered by employing Coq tactics and tacticals. Theorem 1 can then be viewed as a guarantee that verification condition generation and typechecking are terminating processes. This would not be a particularly deep property in any first-order language, but because HTT is higher-order, deciding equality requires *normalization*. This is a non-trivial algorithm, but, by Theorem 1, it terminates.

Theorem 2 (Soundness of the HTT logic) *The judgment $\Delta \vdash P$ cannot derive falsehood, and hence is sound.*

Theorem 2 is established by exhibiting a model of HTT. In [20], we have described a set-theoretic model which takes place in ZFC with infinite inaccessible cardinals $\kappa_0, \kappa_1 \ldots$. All the types are given their obvious set-theoretic interpretation, the universe mono is the set of all sets of cardinality smaller than κ_0, heaps are interpreted as finite functions from nat to $\Sigma\alpha{:}\mathsf{mono}.\,\alpha$, and predicates on a type are interpreted as subsets of the type. Crucially, Hoare types are *interpreted as singleton sets*, and therefore all the computations in HTT are given the same logical interpretation. This is sufficient to argue logical soundness, because HTT currently contains no axioms declaring equations or other kinds of relations on effectful computations (except the monadic laws, which are very simplistic, and are handled by the typechecker). Not surprisingly, the same model suffices to argue the logical soundness of the extension from the current paper.

The above theorems concerns HTT when viewed as a logic. But HTT is at the same time a programming language, and we need to also prove that it is sound when viewed as a programming language. In particular, we need to show that if $I; P \vdash E \Leftarrow x{:}A.\,Q$, then indeed, if E is executed in a shared heap satisfying invariant I and the local heap satisfying the predicate P, and E terminates, then the ending heap satisfies the predicate Q. This theorem would justify our typing rules and show them *adequate* with respect to the intuitive operational interpretation of E.

We will prove this *adequacy theorem* for the current extension of HTT in Section 6, after we have formally defined the operational semantics.

5.2 Identity, consequence, framing, compositionality

Identity principle justifies eta expansion as an equation in the definitional equality of HTT. We prove it here only for the case of Hoare types. For the proofs related to the pure fragment, we refer to [20].

Lemma 3 (Identity principle) *The following statements hold:*

1. *If* $\Delta \vdash K \Rightarrow \mathsf{CMD}\{I\}\{P\}\, x{:}A\{Q\}$, *then* $\Delta \vdash \mathsf{cmdo}\ (x \Leftarrow K; \mathsf{return}\ x) \Leftarrow \mathsf{CMD}\{I\}\{P\}\, x{:}A\{Q\}$.
2. *If* $\Delta \vdash K \Rightarrow \mathsf{ST}\{P\}\, x{:}A\{Q\}$, *then* $\Delta \vdash \mathsf{stdo}\ (x \Leftarrow K; \mathsf{return}\ x) \Leftarrow \mathsf{ST}\{P\}\, x{:}A\{Q\}$.

Proof: We sketch the proof of the first statement. The second one is analogous, but strictly simpler, as there is no need to keep track of the invariant I. To prove the first statement, the typing rules force us to discharge the following verification conditions.

- $\Delta, m{:}\mathsf{heap}, I \vdash I * (I \mathrel{-\!\!*} I)$
- $\Delta, m{:}\mathsf{heap}, P \vdash P * \top$
- $\Delta, x{:}B, \mathsf{i}{:}\mathsf{heap}, \mathsf{m}{:}\mathsf{heap}, \delta P \circ (\exists y{:}B.\, (P \multimap Q) \circ (x = y \wedge \mathsf{i} = \mathsf{m})) \vdash Q$.

But all of these are easy to establish, once the definitions of the propositional connectives are expanded. (We have also proved these properties formally in Coq). ■

Lemma 4 (Principle of consequence) *Suppose that* $\Delta; I; P \vdash E \Leftarrow x{:}A.\, Q$. *Then the following holds:*

1. *Weakening consequent. If* $\Delta, x{:}A, i{:}\mathsf{heap}, m{:}\mathsf{heap}, Q \vdash Q'$, *then* $\Delta; I; P \vdash E \Leftarrow x{:}A.\, Q'$.
2. *Strengthening precedent. If* $\Delta, m{:}\mathsf{heap}, P' \vdash P$, *then* $\Delta; I; P' \vdash E \Leftarrow x{:}A.\, \delta P' \circ Q$.

Similarly for sequential computations. If $\Delta; P \vdash E \Leftarrow x{:}A.\, Q$, *then*

3. *Weakening consequent. If* $\Delta, x{:}A, i{:}\mathsf{heap}, m{:}\mathsf{heap}, Q \vdash Q'$, *then* $\Delta; P \vdash E \Leftarrow x{:}A.\, Q'$.
4. *Strengthening precedent. If* $\Delta, m{:}\mathsf{heap}, P' \vdash P$, *then* $\Delta; P' \vdash E \Leftarrow x{:}A.\, \delta P' \circ Q$.

Proof: Weakening consequent is trivial. By construction, we know that $E \Rightarrow x{:}A.\, T'$ and $\delta P \circ T' \vdash Q$. By assumption, $Q \vdash Q'$ so by cut, $\delta P \circ T' \vdash Q'$ as well. Thus, trivially $E \Leftarrow x{:}A.\, Q'$.

Strengthening precedent is slightly more involved. From the typing of E, we know that $\Delta; I; P \vdash E \Rightarrow x{:}A.\, T'$, for some T' such that

$$\delta P \circ T' \vdash Q.$$

We further observe that

$$\Delta; I; P' \vdash E \Rightarrow x{:}A.\, T'$$

since the generated postcondition never depends on the starting precondition (easily shown by induction). From the first judgment

$$\delta P' \circ \delta P \circ T' \vdash \delta P' \circ Q$$

and since $\delta P' \circ \delta P = \delta P'$, we get

$$\delta P' \circ T' \vdash \delta P' \circ Q.$$

Together with the second judgment, this suffices to conclude the required

$$\Delta; I; P' \vdash E \Leftarrow x{:}A.\, \delta P' \circ Q.$$

■

Lemma 5 (Frame principle) *Suppose that* $\Delta; I; P \vdash E \Leftarrow x{:}A.\, Q$. *Then the following holds:*

1. *Local frame.* $\Delta; I; P * \top \vdash E \Leftarrow x{:}A.\, P \multimap Q$.
2. *Global frame. If* J *is precise, then* $\Delta; I * J; P \vdash E \Leftarrow x{:}A.\, Q$.

Similarly for sequential computations. If $\Delta; P \vdash E \Leftarrow x{:}A.\, Q$, *then*

3. *Frame.* $\Delta; P * \top \vdash E \Leftarrow x{:}A.\, P \multimap Q$.

The presented phrasing of the frame principles may be a bit unusual, when compared to other works on Separation Logic. For example, a more recognizable form of of the local frame principle may be

$$\Delta; I; P * R \vdash E \Leftarrow x{:}A.\, Q \mathbin{**} \delta R$$

which directly states that the initial heap may be extended with an arbitrary subheap satisfying R as long as the ending heap is extended with the same subheap, also satisfying R.

We note that the later form of the frame principle is easily derivable from Lemma 5. Indeed if $\Delta; I; P * \top \vdash E \Leftarrow x{:}A.\, P \multimap Q$, then by strengthening precedent we first get

$$\Delta; I; P * R \vdash E \Leftarrow x{:}A.\, \delta(P * R) \circ (P \multimap Q)$$

and then because $\delta(P * R) \circ (P \multimap Q) \vdash Q \mathbin{**} \delta R$ we can weaken the consequent into

$$\Delta; I; P * R \vdash E \Leftarrow x{:}A.\, Q \mathbin{**} \delta R$$

In order to prove the Frame principle, we first need some auxiliary definitions and lemmas.

Definition 6 (Framing) *Unary predicate* P frames *the binary predicate* T *if, for any heaps* i *and* m *such that* T *holds, we have* $\mathsf{splits}(\mathsf{i}, i_0, h)$ *and* $P\ i_0$ *implies that there exists* m_0 *such that* $\mathsf{splits}(\mathsf{m}, m_0, h)$ *and* $T\ i_0\ m_0$.

Lemma 7 (Associativity of framing) *The following statements hold of arbitrary predicates:*

1. *If* R *does not depend on free heap variables (i.e.,* R *is pure), then* P *frames* δR.
2. *If* $P \vdash P_1 * \top$ *and* $P \circ (P_1 \multimap Q_1)$ *frames* T, *then* P *frames* $(P_1 \multimap Q_1) \circ T$.
3. *If* $P \vdash P_1 * P_2 * \top$ *and* $P \circ (P_1\, P_2 \multimap\!\!\circ\, Q_1\, Q_2)$ *frames* T *then* P *frames* $(P_1\, P_2 \multimap\!\!\circ\, Q_1\, Q_2) \circ T$.
4. *If* P *frames* Q *then* P *frames* $\exists x{:}A.\, Q$.

Proof: Straightforward chasing of subset relations. Formally proved in Coq. ■

We now go back to the Local frame principle (Lemma 5.1), where from the typing of E, we know that $\Delta; I; P \vdash E \Rightarrow x{:}A.\, T$ such that $\delta P \circ T \vdash Q$. To prove this lemma, we first make two auxiliary observations. The first is that

$$\Delta; I; P * R \vdash E \Rightarrow x{:}A.\, T$$

for any predicate R. This property holds because the shape of T depends only on E, but not on the precondition with which the typechecking starts. The precondition only factors into the verification conditions that are required by each of the typing rules, but these verification conditions remain valid after spatially extending the precondition, as conditions to be proved are invariant with respect to the size of the heap (in the usual terminology of Separation Logic, we say that the propositions are *intuitionistic*).

The second observation is that P frames T, in the sense of the definition above. This property is easily proved by induction on the structure of E. In the base case, when $E = \mathsf{return}\ M$, we have that $T = \delta(x = M)$, where $x = M$ is obviously independent of any heap variables. In this case, the result holds by Lemma 7.1.

In the other cases, the result follows from induction hypothesis, also by appropriate cases of Lemma 7, since the strongest postconditions are generated in all the cases by relational composition to the postcondition obtained in the previous step.

Next, if P frames T, then it can easily be shown that

$$\Delta, x{:}A, \mathsf{i}, \mathsf{m}{:}\mathsf{heap}, T \vdash P \multimap Q.$$

The proof basically amounts to the trivial unrolling the definitions of $\multimap$ and framing, and the fact that $\delta P \circ T \vdash Q$. We are now ready to assemble the obtained facts into the proof of Local frame principle. First, we slightly extend the above equation by prefixing T on the left with $\delta(P * R)$, as follows.

$$\Delta, x{:}A, \mathsf{i}, \mathsf{m}{:}\mathsf{heap}, \delta(P * R) \circ T \vdash P \multimap Q.$$

This is obviously valid, as the prefixing just adds new facts about the initial heap before proving entailment. This amounts to weakening and is, of course, admissible. But the above judgment is precisely the premise for deriving

$$\Delta; I; P * R \vdash E \Leftarrow x{:}A.\, P \multimap Q$$

which degenerates to the frame principle if we pick $R = \top$. This concludes the proof of local frame. ■

The proof of the global frame principle is also by induction on the typing of E, knowing that $\Delta; I; P \vdash E \Rightarrow x{:}A.\, T$ where $\delta P \circ T \vdash Q$. As before, it is easy to see that the same T is generated no matter how the invariant I is extended. In other words,

$$\Delta; I * J; P \vdash E \Rightarrow x{:}A.\, T$$

The extension of the invariant does not influence the shape of T, but does influence the verification conditions that are emitted during the checking of E.

There are only two commands where the verification conditions depend on I: monadic bind and atomic execution. The other commands, including $\mathsf{publish}$ (somewhat surprisingly, at first sight), do not generate conditions dependent on I.

In the case $E = y \leftarrow K; E_1$, we know $K \Rightarrow \mathsf{CMD}\,\{I_1\}\{P_1\}\,y{:}B\{Q_1\}$, and by the typing of E:

$$I \vdash I_1 * (I_1 \mathrel{-\!\!*} I)$$

After the extension with J, we will need to show

$$I * J \vdash I_1 * (I_1 \mathrel{-\!\!*} (I * J)).$$

But this immediately follows from the previous entailment, following the definitions of separating connectives.

In the case $E = y \Leftarrow \mathsf{atomic}_{B,R,Q_1}\ E_1; E_2$, we know

$$I * R \vdash E_1 \Leftarrow y{:}B.\, \Box I \mathrel{*\!\!*} Q$$

and we need to show:

$$I * J * R \vdash E_1 \Leftarrow y{:}B.\, \Box(I * R) \mathrel{*\!\!*} Q.$$

To establish this, we first use the sequential frame principle to derive

$$I * R * J \vdash E_1 \Leftarrow y{:}B.\, \Box I \mathrel{*\!\!*} Q \mathrel{*\!\!*} \delta R.$$

Then the weakening of consequent via the entailment $\Box I \mathrel{*\!\!*} Q * \delta R \vdash \Box(I * R) \mathrel{*\!\!*} Q$ produces the result. The entailment holds because $\delta R \vdash \Box R$, trivially. ■

We omit the proof of the frame principle for the sequential judgment, as it is analogous to, but strictly simpler from the Local frame principle.

We further show that HTT is compositional in the sense that typechecking of a program (which amounts to verification) requires only that the individual sub-programs are type-checked separately. There is no need for whole-program reasoning, as the types are strong enough to isolate the program components and serve as their interfaces.

As in any other type theory, HTT's compositionality theorem takes the form of a substitution principle, and we present some selected statements.

Lemma 8 (Substitution principle for terms) *Suppose that* $\Delta \vdash M \Leftarrow A$*, and abbreviate* $[M{:}A/x]T$ *with* T'*, for arbitrary* T*. Then the following holds:*

1. *If* $\Delta, x{:}A, \Delta_1 \vdash K \Rightarrow B$ *then* $\Delta, \Delta_1' \vdash K' \Rightarrow B'$*.*
2. *If* $\Delta, x{:}A, \Delta_1 \vdash N \Leftarrow B$ *then* $\Delta, \Delta_1' \vdash N' \Leftarrow B'$*.*
3. *If* $\Delta, x{:}A, \Delta_1 \vdash P$*, then* $\Delta, \Delta' \vdash P'$*.*
4. *If* $\Delta, x{:}A, \Delta_1; I; P \vdash E \Rightarrow y{:}B.\, Q$ *and* $y \notin \mathsf{FV}(M)$*, then* $\Delta, \Delta_1'; I'; P' \vdash E' \Rightarrow y{:}B'.\, Q'$*.*
5. *If* $\Delta, x{:}A, \Delta_1; I; P \vdash E \Leftarrow y{:}B.\, Q$ *and* $y \notin \mathsf{FV}(M)$*, then* $\Delta, \Delta_1'; I'; P' \vdash E' \Leftarrow y{:}B'.\, Q'$*.*

Similar statements hold for the sequential computations.

6. *If* $\Delta, x{:}A, \Delta_1; P \vdash E \Rightarrow y{:}B.\, Q$ *and* $y \notin \mathsf{FV}(M)$*, then* $\Delta, \Delta_1'; P' \vdash E' \Rightarrow y{:}B'.\, Q'$*.*
7. *If* $\Delta, x{:}A, \Delta_1; P \vdash E \Leftarrow y{:}B.\, Q$ *and* $y \notin \mathsf{FV}(M)$*, then* $\Delta, \Delta_1'; P' \vdash E' \Leftarrow y{:}B'.\, Q'$*.*

Proof: The first three statements are about the pure fragment, so we omit their proofs here – they proceed along the lines presented in previous papers. Here we only check the monadic judgments, as their formulation is new. We start with statement 3.

case $E = \mathsf{return}\ M$. Then $B = A$ and $M \Leftarrow A$ and $Q = \delta(y = M)$. By induction on M, $M' \Leftarrow B'$ and we can return $Q' = \delta(y = M')$.

case $E = x_1 \leftarrow K; E_1$. Here $K \Rightarrow \mathsf{CMD}\,\{I_1\}\{P_1\}\, x_1{:}A_1\{Q_1\}$ and $I \vdash I_1 * (I_1 \mathrel{-\!\!*} I)$ and $P \vdash P_1 * \top$ and $\Delta, x{:}A, \Delta_1, x_1{:}A_1; P \circ (P_1 \multimap Q_1) \vdash E_1 \Rightarrow y{:}B.\, T$ where $Q = (\exists x_1{:}A_1.\, (P_1 \multimap Q_1) \circ T)$. Then by induction on K, E_1 and the derivations of $I \vdash I_1 * (I_1 \mathrel{-\!\!*} I)$ and $P \vdash P_1 * \top$, we get the required.

case $E = (x_1{:}A_1 \Leftarrow E_1{:}P_1 \parallel x_2{:}A_2 \Leftarrow E_2{:}P_2); E_3$. Here too, all the premises directly admit the induction principles.

The rest of the cases are straightforward too. This is to be expected, as typechecking is completely syntax directed. ■

Next we show the main compositionality property for the monadic fragment. It is formulated as a substitution principle about the operation of monadic substitution $\langle E_1/x{:}A\rangle E_2\rangle$. The statement of this substitution principle is essentially and adaptation to binary postconditions of the Hoare-style inference rule for sequential composition. This is an additional aspect of the connection between monadic bind and sequential composition that we mentioned in Section 4.

Lemma 9 (Monadic substitution principle)

1. *If* $\Delta; I; P \vdash E_1 \Rightarrow x{:}A.\, Q$ *and* $\Delta, x{:}A; I; P \circ Q \vdash E_2 \Rightarrow y{:}B.\, T$*, and* $x \notin \mathsf{FV}(B)$ *then* $\Delta; I; P \vdash \langle E_1/x\rangle E_2 \Leftarrow y{:}B.\, (\exists x{:}A.\, Q \circ T)$*.*
2. *If* $\Delta; P \vdash E_1 \Rightarrow x{:}A.\, Q$ *and* $\Delta, x{:}A; P \circ Q \vdash E_2 \Rightarrow y{:}B.\, T$*, and* $x \notin \mathsf{FV}(B)$ *then* $\Delta; P \vdash \langle E_1/x\rangle E_2 \Leftarrow y{:}B.\, (\exists x{:}A.\, Q \circ T)$*.*

Proof: By induction on the structure of E_1.

In case $E_1 = \mathsf{return}\ M$, we have $Q = \delta(x = M)$ and

$$\Delta, x{:}A; I; P \circ \delta(x = M) \vdash E_2 \Rightarrow y{:}B.\, T.$$

Substituting M for x gives

$$\Delta; I; P \circ \delta(M = M) \vdash [M/x]E_2 \Rightarrow y{:}B.\, [M/x]T.$$

Now result follows by consequence, because $P \vdash P \circ \delta(M = M)$ and $[M/x]T \vdash \exists x{:}A.\, \delta(x = M) \circ T$.

case $E_1 = x_1 \leftarrow K; F$. Then we have $\Delta \vdash K \Rightarrow \mathsf{CMD}\,\{I_1\}\{P_1\}\,x_1{:}A_1\{Q_1\}$ and $I \vdash I_1 * (I_1 \mathrel{-\!\!*} I)$ and $P \vdash P_1 * \top$ and $\Delta, x_1{:}A_1; P \circ (P_1 \multimap Q_1) \vdash F \Rightarrow x{:}A.\, T'$, where $Q = \exists x_1{:}A_1.\, (P_1 \multimap Q_1) \circ T'$. From the last equation we easily derive

$$\Delta, x_1{:}A_1, P \circ (P_1 \multimap Q_1) \circ T' \vdash P \circ Q.$$

and by consequence from the typing of E_2,

$$\Delta, x_1{:}A_1; I; P \circ (P_1 \multimap Q_1) \circ T' \vdash E_2 \Rightarrow y{:}B.\, T.$$

By induction hypothesis on F then

$$\Delta, x_1{:}A_1; P \circ (P_1 \multimap Q_1) \vdash \langle F/x\rangle E_2 \Leftarrow y{:}B.\, (\exists x{:}A.\, T' \circ T)$$

and therefore

$$\Delta, x_1{:}A_1; P \circ (P_1 \multimap Q_1) \vdash \langle F/x\rangle E_2 \Rightarrow y{:}B.\, T''$$

where $\delta(P \circ (P_1 \multimap Q_1)) \circ T'' \vdash (\exists x{:}A.\, T' \circ T)$. This last entailment is parametric in i and m, and we will use it below after instantiating these heap variables with appropriate values.

Applying the typing rule for bind to the derivation of $\langle F/x\rangle E_2$, we get

$$\Delta; P \vdash x_1 \leftarrow K; \langle F/x\rangle E_2 \Rightarrow y{:}B.\, (\exists x_1{:}A_1.\, (P_1 \multimap Q_1) \circ T'')$$

It remains to show that

$$\delta P \circ (\exists x_1{:}A_1.\, (P_1 \multimap Q_1) \circ T'') \vdash \exists x{:}A.\, Q \circ T$$

that is, after substituting $Q = \exists x_1{:}A_1.\, (P_1 \multimap Q_1) \circ T'$:

$$\delta P \circ (\exists x_1{:}A_1.\, (P_1 \multimap Q_1) \circ T'') \vdash \exists x{:}A.\, \exists x_1{:}A_1.\, (P_1 \multimap Q_1) \circ T' \circ T$$

To show this, it suffices to show

$$x_1{:}A_1, P\ \mathsf{i}, (P_1 \multimap Q_1)\ i\ h, T''\ h\ m \vdash \exists x{:}A.\, ((P_1 \multimap Q_1) \circ T' \circ T)\ i\ m$$

But, the left side implies $P\ \mathsf{i}, (P_1 \multimap Q_1)\ i\ h, (\delta(P \circ (P_1 \multimap Q_1)) \circ T'')\ h\ m$, and thus also $P\ \mathsf{i}, (P_1 \multimap Q_1)\ i\ h, \exists x{:}A.\, (T' \circ T)\ h\ m$. Therefore, it suffices to show

$$x_1{:}A_1, P\ \mathsf{i}, (P_1 \multimap Q_1)\ i\ h, \exists x{:}A.\, (T' \circ T)\ h\ m \vdash \exists x{:}A.\, ((P_1 \multimap Q_1) \circ T' \circ T)\ i\ m$$

But the later is obviously true, by definition of composition.

case $E = (x_1{:}A_1 \Leftarrow E_1{:}P_1 \parallel x_2{:}A_2 \Leftarrow E_2{:}P_2); F$. By the typing rules we know

$$\Delta, x_1{:}A_1, x_2{:}A_2; I; P \circ (P_1\, P_2 \multimap\!\!\circ\, Q_1\, Q_2) \vdash F \Rightarrow T'$$

where $Q = \exists x_1{:}A_1.\, \exists x_2{:}A_2.\, (P_1\, P_2 \multimap\!\!\circ\, Q_1\, Q_2) \circ T'$, and thus

$$x_1{:}A_1, x_2{:}A_2, P \circ (P_1\, P_2 \multimap\!\!\circ\, Q_1\, Q_2) \circ T' \vdash P \circ Q$$

By consequence principle applied to the typing of E_2:

$$\Delta, x{:}A, x_1{:}A_1, x_2{:}A_2; I; P \circ (P_1\ P_2 \multimap\!\!\circ\ Q_1\ Q_2) \circ T' \vdash E_2 \Rightarrow y{:}B.\, T$$

and by induction on F:

$$\Delta, x_1{:}A_1, x_2{:}A_2; I; P \circ (P_1\ P_2 \multimap\!\!\circ\ Q_1\ Q_2) \vdash \langle F/x\rangle E_2 \Leftarrow y{:}B.\, (\exists x{:}A.\, T' \circ T)$$

Correspondingly, there exists T'' such that

$$\Delta, x_1{:}A_1, x_2{:}A_2; I; P \circ (P_1\ P_2 \multimap\!\!\circ\ Q_1\ Q_2) \vdash \langle F/x\rangle E_2 \Rightarrow y{:}B.\, T''$$

where $\delta(P \circ (P_1\ P_2 \multimap\!\!\circ\ Q_1\ Q_2)) \circ T'' \vdash \exists x{:}A.\, T' \circ T$.

Applying the typing rule for parallelism to the derivation of $\langle F/x\rangle E_2$, we get

$$\Delta; P \vdash (x_1{:}A_1 \Leftarrow E_1{:}P_1 \parallel x_2{:}A_2 \Leftarrow E_2{:}P_2); \langle F/x\rangle E_2 \Rightarrow y{:}B.\, (\exists x_1{:}A_1.\, \exists x_2{:}A_2.\, (P_1\ P_2 \multimap\!\!\circ\ Q_1\ Q_2) \circ T'')$$

It remains to show that

$$\delta P \circ (\exists x_1{:}A_1.\, \exists x_2{:}A_2.\, (P_1\ P_2 \multimap\!\!\circ\ Q_1\ Q_2) \circ T'') \vdash \exists x{:}A.\, Q \circ T$$

that is, after substituting $Q = \exists x_1{:}A_1.\, \exists x_2{:}A_2.\, (P_1\ P_2 \multimap\!\!\circ\ Q_1\ Q_2) \circ T'$:

$$\delta P \circ (\exists x_1{:}A_1.\, \exists x_2{:}A_2.\, (P_1\ P_2 \multimap\!\!\circ\ Q_1\ Q_2) \circ T'') \vdash \exists x{:}A.\, \exists x_1{:}A_1.\, \exists x_2{:}A_2.\, (P_1\ P_2 \multimap\!\!\circ\ Q_1\ Q_2) \circ T' \circ T$$

This is easily shown, following the proof of the previous case.

case $E = x_1 \Leftarrow \mathsf{atomic}_{A_1,P_1,Q_1}\ E_1; F$. By the typing rules, we know:

$$\Delta, x_1{:}A_1; I; P \circ (P_1 \multimap Q_1) \vdash F \Rightarrow x{:}A.\, T'$$

where $Q = \exists x_1{:}A_1.\, (P_1 \multimap Q_1) \circ T'$, and thus

$$x_1{:}A_1, P \circ (P_1 \multimap Q_1) \circ T' \vdash P \circ Q$$

By consequence applied to typing of E_2:

$$\Delta, x{:}A, x_1{:}A_1; I; P \circ (P_1 \multimap Q_1) \circ T' \vdash E_2 \Rightarrow y{:}B.\, T$$

and by induction on F:

$$\Delta, x_1{:}A_1; I; P \circ (P_1 \multimap Q_1) \circ T' \vdash \langle F/x\rangle E_2 \Leftarrow y{:}B.\, (\exists x{:}A.\, T' \circ T).$$

Therefore

$$\Delta, x_1{:}A_1; I; P \circ (P_1 \multimap Q_1) \circ T' \vdash \langle F/x\rangle E_2 \Rightarrow y{:}B.\, T''$$

where $\delta(P \circ (P_1 \multimap Q_1)) \circ T'' \vdash \exists x{:}A.\, T' \circ T$. Then the proof is finished off in the completely analogous way to the case for monadic bind.

case $E = \mathsf{publish} J; F$, where

$$\Delta; I * J; P \circ (J \multimap \nabla\mathsf{emp}) \vdash F \Rightarrow x{:}A.\, T'$$

where $Q = (J \multimap \nabla\mathsf{emp}) \circ T'$, and thus

$$P \circ (J \multimap \nabla\mathsf{emp}) \circ T' \vdash P \circ Q$$

Then we proceed as in the previous cases, to first apply consequence to E_2, induction on F, then prove a customary entailment before putting the things back together. The point is, the entailment is always true because the typing rules are organized so that the postcondition obtained in one step is always pre-composed with another relation to compute the new postcondition.

The proofs of the other cases also follow this pattern.

■

5.3 Inversion

With an eye towards operational semantics, in this section we capture the process of evolving the precondition of a computation as it steps through its execution. We will formally define the operational semantics in the next section; here we develop several background notions and lemmas.

First, we need to model the process of selecting the expression, among a set of expressions appearing in parallel, whose evaluation will be advanced in the next operational step. We require a new syntactic domain of continuations κ, defined as follows.

$$\begin{array}{lrcl} \textit{Continuations} & \kappa\,[P,E] & ::= & (x_1{:}A_1 \Leftarrow E_1{:}P_1 \parallel x_1{:}A_2 \Leftarrow E{:}P); E_3 \mid \\ & & & (x_1{:}A_1 \Leftarrow E{:}P \parallel x_2{:}A_2 \Leftarrow E_2{:}P_2); E_3 \end{array}$$

Given a list of continuations $\overline{\kappa_i} = (\kappa_1, \ldots, \kappa_n)$, and a list of predicates $\overline{P_i} = (P_1, \ldots, P_n)$, and an expression E, we write

$$\overline{\kappa_i}\,\overline{P_i}\,E \quad = \quad \kappa_1[P_1, \cdots \kappa_{n-1}[P_{n-1}, \kappa_n[P_n, E]]]$$

Thus, the list $\overline{\kappa_i}$ determines the sequence of parallel nestings at the bottom of which E appears, while the list $\overline{P_i}$ determines the sequence of footprints provided along the path.

Lemma 10 (Inversion) *The following are true:*

1. *If* $\Delta; I; P \vdash \overline{\kappa_i}\,\overline{P_i}\,(y \Leftarrow \mathsf{atomic}_{B,R,Q}\,E; E_1) \Leftarrow x{:}A.\,T$, *then* $\Delta, y{:}B; I; P \circ (R \multimap Q) \vdash \overline{\kappa_i}\,\overline{(P_i \circ (R \multimap Q))}\,E_1 \Rightarrow x{:}A.\,T'$ *and* $y{:}B, i{:}\mathsf{heap}, m{:}\mathsf{heap}, \delta P \circ (R \multimap Q) \circ T' \vdash T$.
2. *If* $\Delta; I; P \vdash \overline{\kappa_i}\ \overline{P_i}\ (\mathsf{publish}\ J; E) \Leftarrow x{:}A.\,T$, *then* $\Delta; I * J; P \circ (J \multimap \nabla\mathsf{emp}) \vdash \overline{\kappa_i}\ \overline{(P_i \circ (J \multimap \nabla\mathsf{emp}))}\ E \Rightarrow x{:}A.\,T'$ *and* $i{:}\mathsf{heap}, m{:}\mathsf{heap}, \delta P \circ (J \multimap \nabla\mathsf{emp}) \circ T' \vdash T$.

Proof: The statements about the state monad are trivially proved by inversion on the typing of the involved expressions. The statements about the command monad are more involved. We only attempt statements 1, 2 and 3. Statement 1 is trivial, just by unraveling the definition of the application $\bar{\kappa}\,\bar{P}\,E$. Statement 2 is by induction on n. In case $n = 0$, then $\Delta; I; P \vdash y \Leftarrow \mathsf{atomic}_{B,R,Q}\,E; E_1 \Leftarrow x{:}A.\,T$, and therefore

$$\Delta; I; P \vdash y \Leftarrow \mathsf{atomic}_{B,R,Q}\,E; E_1 \Rightarrow x{:}A.\,Q'$$

for some Q' such that $\delta P \circ Q' \vdash T$. One of the premises of the typing rule for atomic is:

$$\Delta, y{:}B; I; P \circ (R \multimap Q) \vdash E_1 \Rightarrow x{:}A.\,T'$$

where $Q' = \exists y{:}B.\,(R \multimap Q) \circ T'$. Then the result trivially follows.

In case $n > 0$. We have $\Delta; I; P \vdash \kappa_1\,P_1\,(\overline{\kappa_i}\,\overline{P_i}\,(y \Leftarrow \mathsf{atomic}_{B,R,Q}\,E; E_1)) \Leftarrow x{:}A.\,T$, and therefore,

$$\Delta; I; P \vdash \kappa_1\,P_1\,(\overline{\kappa_i}\,\overline{P_i}\,(y \Leftarrow \mathsf{atomic}_{B,R,Q}\,E; E_1)) \Rightarrow x{:}A.\,Q'$$

for some Q' such that $\delta P \circ Q' \vdash T$. We also know that $P \vdash R * \top$, and $P_1 \vdash R * \top$ and $P_i \vdash R * \top$. Without loss of generality, we assume that $\kappa_1 = x_1{:}A_1 \Leftarrow -{:}- \parallel x_2{:}A_2 \Leftarrow E_2{:}P_2; E_3$.

From this typing, we can infer:

$$\Delta; I; P_1 \vdash \overline{\kappa_i}\,\overline{P_i}\,(y \Leftarrow \mathsf{atomic}_{B,R,Q}\,E; E_1) \Rightarrow x_1{:}A_1.\,Q_1$$

and

$$\Delta; I; P_2 \vdash E_2 \Rightarrow x_2{:}A_2.\,Q_2$$

and

$$P \vdash P_1 * P_2 * \top$$

and

$$\Delta, x_1{:}A_1, x_2{:}A_2; I; P \circ (P_1\ P_2 \multimap\!\circ\ Q_1\ Q_2) \vdash E_3 \Rightarrow x{:}A.\, Q''$$

where $Q' = \exists x_1{:}A_1.\, \exists x_2{:}A_2.\, (P_1\ P_2 \multimap\!\circ\ Q_1\ Q_2) \circ Q''$.

From the first of these four judgments, by induction hypothesis:

$$\Delta, y{:}B; I; P_1 \circ (R \multimap Q) \vdash \overline{\kappa_i}\ \overline{(P_i \circ (R \multimap Q))}\ E_1 \Rightarrow x_1{:}A_1.\, T''$$

where $\delta P_1 \circ (R \multimap Q) \circ T'' \vdash Q_1$.

By weakening, we first introduce y into the typing of E_3.

$$\Delta, x_1{:}A_1, x_2{:}A_2, y{:}B; I; P \circ (P_1\ P_2 \multimap\!\circ\ Q_1\ Q_2) \vdash E_3 \Rightarrow x{:}A.\, Q''$$

Then we notice that the following is true[3]:

$$y{:}B; P \circ (R \multimap Q) \circ [(P_1 \circ (R \multimap Q))\ (P_2) \multimap\!\circ\ (T'')\ (Q_2)] \vdash P \circ (P_1\ P_2 \multimap\!\circ\ Q_1\ Q_2)$$

Thus, by consequence:

$$\Delta, x_1{:}A_1, x_2{:}A_2, y{:}B; I; P \circ (R \multimap Q) \circ [(P_1 \circ (R \multimap Q))\ (P_2) \multimap\!\circ\ (T'')\ (Q_2)] \vdash E_3 \Rightarrow x{:}A.\, Q''$$

We also notice (also proved in Coq) that:

$$P \circ (R \multimap Q) \vdash (P_1 \circ (R \multimap Q)) * P_2 * \top$$

and therefore we can put the components together, to conclude:

$$\begin{aligned}\Delta, y{:}B; I; P \circ (R \multimap Q) \vdash \kappa_1\ (P_1 \circ (R \multimap Q))\ (\overline{\kappa_i}\ \overline{(P_i \circ (R \multimap Q))}\ E_1) \\ \Rightarrow x{:}A.\, \exists x_1{:}A_1.\, \exists x_2{:}A_2.\, [(P_1 \circ (R \multimap Q))\ (P_2) \multimap\!\circ\ (T'')\ (Q_2)] \circ Q''.\end{aligned}$$

But now, $\delta P \circ (R \multimap Q) \circ [\multimap\!\circ\ (P_1 \circ (R \multimap Q))(P2)(T'')(Q_2)] \circ Q''$ implies $\delta P \circ (P_1\ P_2 \multimap\!\circ\ Q_1\ Q_2) \circ Q''$ by the above equation. Then, taking into account the equality of Q', we get that the later equals $\delta P \circ Q'$. And we know that this implies T. This concludes the proof.

Statement 2 is also proved by induction on n.

In case $n = 0$, we have $\Delta; I; P \vdash \mathsf{publish}\ J; E_1 \Rightarrow x{:}A.\, Q'$, where $\delta P \circ Q' \vdash T$. By inversion on the typing rule, we know $P \vdash J * \top$ and

$$\Delta; I * J; P \circ (J \multimap \nabla\mathsf{emp}) \vdash E_1 \Rightarrow x{:}A.\, T'$$

where $Q' = (J \multimap \nabla\mathsf{emp}) \circ T'$. This immediately proves the case.

In case $n > 0$, we have $\Delta; I; P \vdash \kappa_1\ P_1\ (\overline{\kappa_i}\ \overline{P_i}(\mathsf{publish}\ J; E_1)) \Leftarrow x{:}A.\, T$, and therefore

$$\Delta; I; P \vdash \kappa_1\ P_1\ \overline{\kappa_i}\ \overline{P_i}(\mathsf{publish}\ J; E_1)) \Rightarrow x{:}A.\, Q'$$

for some Q' such that $\delta P \circ Q' \vdash T$. We also know $P \vdash J * \top$ and $P_1 \vdash J * \top$ and $P_i \vdash J * \top$. Wlog, assume that $\kappa_1 = x_1{:}A_1 \Leftarrow -{:}- \parallel x_2{:}A_2 \Leftarrow E_2{:}P_2; E_3$.

From this typing, we can infer:

$$\Delta; I; P_1 \vdash \overline{\kappa_i}\,\overline{P_i}\ (\mathsf{publish}\ J; E_1) \Rightarrow x_1{:}A_1.\, Q_1$$

and

$$\Delta; I; P_2 \vdash E_2 \Rightarrow x_2{:}A_2.\, Q_2$$

[3]and proved in Coq

and

$$P \vdash P_1 * P_2 * \top$$

and

$$\Delta, x_1{:}A_1, x_2{:}A_2; I; P \circ (P_1\, P_2 \multimap\!\circ\, Q_1\, Q_2) \vdash E_3 \Rightarrow x{:}A.\, Q''$$

where $Q' = \exists x_1{:}A_1.\, \exists x_2{:}A_2.\, (P_1\, P_2 \multimap\!\circ\, Q_1\, Q_2) \circ Q''$.

From the first of these four judgments, by induction hypothesis:

$$\Delta; I * J; P_1 \circ (J \multimap \nabla\mathsf{emp}) \vdash \overline{\kappa_i}\, \overline{(P_i \circ (J \multimap \nabla\mathsf{emp}))}\, E \Rightarrow x_1{:}A_1.\, T''$$

where $\delta P \circ (J \multimap \nabla\mathsf{emp}) \circ T'' \vdash Q_1$.

Then we notice that the following is true, as a special instance of the theorem proved for the previous case, by taking $R = J$ and $Q = \nabla\mathsf{emp}$:

$$P \circ (J \multimap \nabla\mathsf{emp}) \circ ((P_1 \circ (J \multimap \nabla\mathsf{emp}))\, (P_2) \multimap\!\circ\, (T'')\, (Q_2)) \vdash P \circ (P_1\, P_2 \multimap\!\circ\, Q_1\, Q_2)$$

Thus, by consequence:

$$\Delta, x_1{:}A_1, x_2{:}A_2; I * J; P \circ (J \multimap \nabla\mathsf{emp}) \circ ((P_1 \circ (J \multimap \nabla\mathsf{emp}))\, (P_2) \multimap\!\circ\, (T'')\, Q_2) \vdash E_3 \Rightarrow x{:}A.\, Q''$$

We also notice, again as a special instance of the equation from the previous case that:

$$P \circ (J \multimap \nabla\mathsf{emp}) \vdash (P_1 \circ (J \multimap \nabla\mathsf{emp})) * P_2 * \top$$

Furthermore, by the global frame rule:

$$\Delta; I * J; P_2 \vdash E_2 \Rightarrow x_2{:}A_2.\, Q_2$$

and therefore we can put the components together to obtain:

$$\begin{aligned}\Delta; I; P \circ (J \multimap \nabla\mathsf{emp}) \vdash \kappa_1\ (P_1 \circ (J \multimap \nabla\mathsf{emp}))\ (\overline{\kappa_i}\ \overline{(P_i \circ (J \multimap \nabla\mathsf{emp}))}\ E_1)\\ \Rightarrow x{:}A.\, \exists x_1{:}A_1.\, \exists x_2{:}A_2.\, ((P_1 \circ (J \multimap \nabla\mathsf{emp}))\, (P_2) \multimap\!\circ\, (T'')\, (Q_2)) \circ Q''.\end{aligned}$$

But now $\delta P \circ (J \multimap \nabla\mathsf{emp}) \circ [(P_1 \circ (J \multimap \nabla\mathsf{emp}))\, (P_2) \multimap\!\circ\, (T'')\, (Q_2)] \circ Q''$ implies $\delta P \circ (P_1\, P_2 \multimap\!\circ\, Q_1\, Q_2) \circ Q''$, and the later equals $\delta P \circ Q'$ which entails T. Thus taking

$$T' = \exists x_1{:}A_1.\, \exists x_2{:}A_2.\, [(P_1 \circ (J \multimap \nabla\mathsf{emp}))\, (P_2) \multimap\!\circ\, (T'')\, (Q_2)] \circ Q''$$

satisfies the requirement of the lemma.

■

6 Operational semantics

In this section we focus on the operational semantics of the monadic fragment of HTT, and prove theorems about ST and CMD-computations. The purely functional fragment is quite standard. Since the functional fragment is a sub-language of ECC, we know that it is *strongly normalizing*. Therefore, we can give the functional fragment a number of reduction strategies, including call-by-name and call-by-value. Alternatively, we can *normalize* all of the pure subterms before applying the evaluation rules for the monadic terms. Thus, we are justified in blurring the distinction between pure terms and *values* as each pure term evaluates to a unique value (in the usual sense employed in functional programming) equal to its *canonical form*. We omit the treatment of the pure fragment, and refer the interested reader to [20].

The operational semantics of monadic computations requires the following syntactic categories.

$$\begin{array}{lrcl}
\textit{Run-time heaps} & \chi & ::= & \cdot \mid \chi, l \mapsto_\tau M \\
\textit{Abstract machines} & \mu & ::= & (\chi, E) \mid (\chi, E_1 \mid x{:}A.\, E_2) \\
\textit{Stacks} & \kappa\,[P, E] & ::= & (x_1{:}A_1 \Leftarrow E_1{:}P_1 \parallel x_1{:}A_2 \Leftarrow E{:}P); E_3 \mid \\
 & & & (x_1{:}A_1 \Leftarrow E{:}P \parallel x_2{:}A_2 \Leftarrow E_2{:}P_2); E_3
\end{array}$$

Run-time heaps are finite maps from locations to terms. These are the objects about which our assertions logic reasons. The soundness of the assertion logic established in Theorem 2 makes the connection between the run-time behavior of HTT and its logical behavior. If HTT shows that at some point in the program the heap should contain a location l pointing to a value $M{:}\tau$, then, when that point in the program is reached at run-time, the heap contains an assignment $l \mapsto_\tau M$.

Abstract machines pair up a run-time heap with an expression to be evaluated. They come in two modes:

1. (χ, E) is the *concurrent mode*, which takes a CMD expression E describing the concurrent execution of a number of processes; and
2. $(\chi, E_1 \mid x{:}A.\, E_2)$ is the *atomic mode*. In the atomic mode, E_1 is an ST-computation, which must be executed before returning to the (concurrent) continuation E_2. The value of E_1 is bound to the variable $x{:}A$ in E_2.

Stacks are used to select an expression from a set of parallel expressions in E. The selected expression will be advanced one step according to the operational semantics. Given a list of stacks $\overline{\kappa_i} = (\kappa_1, \ldots, \kappa_n)$, and a list of predicates $\overline{P_i} = (P_1, \ldots, P_n)$, and an expression E, we write $\overline{\kappa_i}\,\overline{P_i}\,E$ as an abbreviation for $\kappa_1[P_1, \cdots \kappa_{n-1}[P_{n-1}, \kappa_n[P_n, E]]]$. Thus, the list $\overline{\kappa_i}$ determines the sequence of parallel nestings, at the bottom of which E appears, and the list $\overline{P_i}$ determines the sequence of footprint annotations provided along the path.

The main judgment of the operational semantics has the form $\mu \hookrightarrow \mu'$. We present the main redexes for concurrent configurations.

$$\frac{}{\chi, \overline{\kappa_i}\ \overline{P_i}\ (x \Leftarrow \mathsf{atomic}_{A,R,Q}\, E_1; E) \ \hookrightarrow \chi, E_1 \mid x{:}A.\, \overline{\kappa_i}\ \overline{(P_i \circ (R \multimap Q))}\ E}$$

$$\frac{}{\chi, \overline{\kappa_i}\ \overline{P_i}\ (\mathsf{publish}\ J; E) \hookrightarrow \chi, \overline{\kappa_i}\ \overline{(P_i \circ (J \multimap \nabla\mathsf{emp}))}\ E}$$

$$\frac{}{\chi, \overline{\kappa_i}\ \overline{P_i}\ ((x_1{:}A_1 \Leftarrow \mathsf{return}\ M_1{:}P_1 \parallel x_2{:}A_2 \Leftarrow \mathsf{return}\ M_2{:}P_2); E) \hookrightarrow \chi, \overline{\kappa_i}\ \overline{P_i}\ ([M_1{:}A_1/x_1, M_2{:}A_2/x_2]E)}$$

$$\frac{}{\chi, \overline{\kappa_i}\ \overline{P_i}\ (x_1 \Leftarrow \mathsf{fix}_B\, f; E) \hookrightarrow \chi, \overline{\kappa_i}\ \overline{P_i}\ (x_1 \Leftarrow f(\mathsf{cmdo}\ (y \Leftarrow \mathsf{fix}_B\, f; \mathsf{return}\ y)); E)}$$

$$\frac{}{\chi, \overline{\kappa_i}\ \overline{P_i}\ (x \Leftarrow (\mathsf{cmdo}\ E_1){:}\mathsf{CMD}\,\{I\}\{R_1\}\,x{:}A\{Q_1\}; E) \ \hookrightarrow \chi, \overline{\kappa_i}\ \overline{P_i}\ (\langle E_1/x{:}A\rangle E)}$$

$$\frac{}{\chi, \overline{\kappa_i}\ \overline{P_i}\ (x \Leftarrow \mathsf{if}_A\ \mathsf{tt}\ \mathsf{then}\ E_t\ \mathsf{else}\ E_f; E) \hookrightarrow \chi, \overline{\kappa_i}\ \overline{P_i}\ (\langle E_t/x{:}A\rangle E)}$$

$$\frac{}{\chi, \overline{\kappa_i}\ \overline{P_i}\ (x \Leftarrow \mathsf{if}_A\ \mathsf{ff}\ \mathsf{then}\ E_t\ \mathsf{else}\ E_f; E) \hookrightarrow \chi, \overline{\kappa_i}\ \overline{P_i}\ (\langle E_f/x{:}A\rangle E)}$$

The rules use a list of stacks $\overline{\kappa_i}\ \overline{P_i}$ to select the first command to execute. Many different possibilities may arise corresponding to different selected stacks, reflecting the the non-deterministic nature of concurrent evaluation.

In the case of atomic, once a command E_1 is selected for atomic execution, the abstract machine moves into the atomic configuration, where E_1 proceeds to be executed without interference from other processes, and with exclusive access to the heap χ[4].

[4]The optimistic evaluation usually associated with transactions need not be reflected in the operational semantics: it is an implementation strategy for speeding up the execution that does not change the semantics.

Upon making a step, both atomic and publish change the local heap, and the annotations encountered along the stack list, $\overline{\kappa_i}$, must be updated in order to reflect the new heap values. In the case of atomic, we use the predicate list $\overline{(P_i \circ (R \multimap Q))}$, because the execution of E_1 is captured by the relation $R \multimap Q$. In the case of publish, we use the predicate list $\overline{(P_i \circ (J \multimap \nabla\mathsf{emp}))}$, because the execution of publish must erase the space described by J, and this operation is captured by the relation $J \multimap \nabla\mathsf{emp}$.

The rules for the atomic configurations are straightforward, so we present the characteristic ones without any comments.

$$\overline{(\chi, l \mapsto_\tau M), (x \Leftarrow !_\tau l; E) \mid y{:}A.\, E_1 \hookrightarrow (\chi, l \mapsto_\tau M), [M{:}\tau/x]E \mid y{:}A.\, E_1}$$

$$\overline{\chi, (x \Leftarrow \mathsf{alloc}_\tau M; E) \mid y{:}A.\, E_1 \hookrightarrow (\chi, l \mapsto_\tau M), [l{:}\mathsf{nat}/x]E \mid y{:}A.\, E_1}$$

$$\overline{(\chi, l \mapsto -), (l :=_\tau M; E) \mid y{:}A.\, E_1 \hookrightarrow (\chi, l \mapsto_\tau M), E \mid y{:}A.\, E_1}$$

$$\overline{(\chi, l \mapsto -), (\mathsf{dealloc}\ l; E) \mid y{:}A.\, E_1 \hookrightarrow \chi, E \mid y{:}A.\, E_1}$$

$$\overline{\chi, (x \Leftarrow (\mathsf{stdo}\ E_2){:}\mathsf{ST}\,\{R_1\}\, x{:}B\{Q_1\}; E) \mid y{:}A.\, E_1 \hookrightarrow \chi, \langle E_2/x{:}B\rangle E \mid y{:}A.\, E_1}$$

$$\overline{\chi, (\mathsf{return}\ M) \mid y{:}A.\, E_1 \hookrightarrow \chi, [M{:}A/y]E_1}$$

Next we define a translation from a value heap into a heap predicate.

$$\begin{array}{rcl} [\![\cdot]\!] & = & \mathsf{emp} \\ [\![\chi, l \mapsto_\tau M]\!] & = & [\![\chi]\!] * l \mapsto_\tau M \end{array}$$

We define a typing judgment for abstract machines. The judgment has the form $I \vdash \mu \Leftarrow x{:}A.\, S$ if, informally, the machine μ preserves the heap invariant I and if its execution terminates, then the heap satisfies the predicate S. More formally:

Definition 11 *We say that $I \vdash \mu \Leftarrow x{:}A.\, S$ if*

1. *$\mu = \chi, E$ and $\chi = \chi_1, \chi_2$ such that $[\![\chi_1]\!] \vdash I$ and $I; [\![\chi_2]\!] \vdash E \Leftarrow x{:}A.\, \nabla S$, or*
2. *$\mu = \chi, E_1 \mid y{:}B.\, E_2$ then there exists a predicate R such that $[\![\chi]\!] \vdash E_1 \Leftarrow y{:}B.\, \nabla(I * R)$ and $y{:}B; I; R \vdash E_2 \Leftarrow x{:}A.\, \nabla S$.*

Notice that in the above definition, the case $\mu = \chi, E_1, y{:}B.\, E_2$ does not impose any requirements that the invariant I holds of χ. This is because the system is designed precisely to allow for breaking of this invariant in the expressions executed atomically (like E_1 above), as long as the invariant is restored before the rest of the computation (E_2 above) is scheduled for execution.

Theorem 12 (Preservation) *If $I \vdash \mu \Leftarrow x{:}A.\, S$ and $\mu \hookrightarrow \mu'$, then $I \vdash \mu' \Leftarrow x{:}A.\, S$.*

Proof: By case analysis on μ.

case $\mu = \chi, \overline{\kappa_i}\ \overline{P_i}\ (y \Leftarrow \mathsf{atomic}_{B,R,Q}\ E_1; E_2)$. Then $\chi = \chi_1, \chi_2$ and $[\![\chi_1]\!] \vdash I$ and

$$I; [\![\chi_2]\!] \vdash \overline{\kappa_i}\ \overline{P_i}\ (y \Leftarrow \mathsf{atomic}_{B,R,Q}\ E_1; E_2) \Leftarrow x{:}A.\, \nabla S.$$

From this typing, we also derive:

$$[\![\chi_2]\!] \vdash R * \top$$

and

$$I * R \vdash E_1 \Leftarrow y{:}B.\, \Box I \mathbin{-\!\!*} Q$$

and, by Inversion Lemma (Lemma 10):

$$y{:}B; I; [\![\chi_2]\!] \circ (R \multimap Q) \vdash \overline{\kappa_i}\ \overline{(P_i \circ (R \multimap Q))}\ E_2 \Rightarrow x{:}A.\, T'.$$

where $y{:}B, \delta([\![\chi_2]\!]) \circ (R \multimap Q) \circ T' \vdash \nabla S$. Since ∇S does not depend on i, we can arbitrarily set the value of i on the left side of this judgment. In particular, we have

$$y{:}B, \delta([\![\chi_2]\!] \circ (R \multimap Q)) \circ T' \vdash \nabla S$$

and therefore:

$$y{:}B; I; [\![\chi_2]\!] \circ (R \multimap Q) \vdash \overline{\kappa_i}\ \overline{(P_i \circ (R \multimap Q))}\ E_2 \Leftarrow x{:}A.\, \nabla S.$$

Note that this property doesn't hold if we use a general binary predicate as a postcondition in the judgment for operational semantics. Insisting on ∇ in ∇S is crucial.

Let $\chi_2 = \chi_3, \chi_4$ where $\chi_3 \vdash R$.

From the typing of E_1, by local frame, we get:

$$I * R * \top \vdash E_1 \Leftarrow y{:}B.\, (I * R) \multimap (\Box I \mathbin{**} Q)$$

By refinement property:

$$[\![\chi]\!] \vdash E_1 \Leftarrow y{:}B.\, \delta[\![\chi]\!] \circ (I * R \multimap \Box I \mathbin{**} Q)$$

Now, we can prove easily (done in Coq) that:

$$\delta[\![\chi]\!] \circ (I * R \multimap \Box I \mathbin{**} Q) \vdash \nabla(I * ([\![\chi_2]\!] \circ (R \multimap Q)))$$

and thus by consequence:

$$[\![\chi]\!] \vdash E_1 \Leftarrow y{:}B.\, \nabla(I * ([\![\chi_2]\!] \circ (R \multimap Q)))$$

In combination with the already established

$$y{:}B; I; [\![\chi_2]\!] \circ (R \multimap Q) \vdash \overline{\kappa_i}\ \overline{(P_i \circ (R \multimap Q))}\ E_2 \Leftarrow x{:}A.\, \nabla S.$$

this leads to

$$I \vdash \mu' \Leftarrow x{:}A.\, S$$

where

$$\mu' = \chi, E_1, y{:}B.\, \overline{\kappa_i}\ \overline{(P_i \circ (R \multimap Q))}\ E_2.$$

But this is exactly what was required.

case $\mu = \chi, (z \Leftarrow \mathsf{alloc}_\tau\ M; E_1), y{:}B.\, E_2$. By typing of μ we know

$$[\![\chi]\!] z \Leftarrow \mathsf{alloc}_\tau\ M; E_1 \Leftarrow y{:}B.\, \nabla(I * R)$$

and

$$y{:}B; I; R \vdash E_2 \Leftarrow x{:}A.\, \nabla S$$

By inversion on the typing of allocation,

$$[\![\chi]\!] \circ (\mathsf{emp} \multimap \nabla(z \mapsto_\tau M)) \vdash E_1 \Rightarrow T$$

where $\delta[\![\chi]\!] \circ (\mathsf{emp} \multimap \nabla(z \mapsto_\tau M)) \circ T \vdash \nabla(I * R)$. After substituting z by a fresh l, we get:

$$[\![\chi, l \mapsto_\tau M]\!] \vdash E_1 \Leftarrow \nabla(I * R)$$

and since $\mu' = (\chi, l \mapsto_\tau M), E_1, y{:}B.\, E_2$, we can assemble back that

$$I \vdash \mu' \Leftarrow x{:}A.\, S$$

case $\mu = (\chi, l \mapsto_C w), (l :=_\tau M; E_1), y{:}B.\, E_2$, for some w of an arbitrary type C. By the typing of μ, we know

$$[\![\chi, l \mapsto_C w]\!] \vdash l :=_\tau M; E_1 \Leftarrow y{:}B.\, \nabla(I * R)$$

and

$$y{:}B; I; R \vdash E_2 \Leftarrow x{:}A.\, \nabla S.$$

By the typing of mutation:

$$[\![\chi, l \mapsto_C w]\!] \circ (l \mapsto - \multimap \nabla l \mapsto_\tau M) \vdash E_1 \Rightarrow T$$

where $\delta[\![\chi, l \mapsto_C w]\!] \circ (l \mapsto - \multimap \nabla l \mapsto_\tau M) \circ T \vdash \nabla(I * R)$. Therefore

$$[\![\chi, l \mapsto_\tau M]\!] \vdash E_1 \Leftarrow \nabla(I * R)$$

and since $\mu' = (\chi, l \mapsto_\tau M), E_1, y{:}B.\, E_2$, we can assemble back that

$$I \vdash \mu' \Leftarrow x{:}A.\, S$$

case $\mu = (\chi, l \mapsto_C w), (\mathsf{dealloc}\ l; E_1), z{:}B.\, E_2$. By the typing of μ, we know

$$[\![\chi, l \mapsto_C w]\!] \vdash \mathsf{dealloc}\ l; E_1 \Leftarrow y{:}B.\, \nabla(I * R)$$

and

$$y{:}B; I; R \vdash E_2 \Leftarrow x{:}A.\, \nabla S.$$

By the typing of deallocation

$$[\![\chi, l \mapsto_C w]\!] \circ (l \mapsto - \multimap \nabla \mathsf{emp}) \vdash E_1 \Rightarrow T$$

where $\delta[\![\chi, l \mapsto_C w]\!] \circ (l \mapsto - \multimap \nabla \mathsf{emp}) \circ T \vdash \nabla(I * R)$. Therefore

$$[\![\chi]\!] \vdash E_1 \Leftarrow \nabla(I * R)$$

and since $\mu' = \chi, E_1, y{:}B.\, E_2$, we can assemble back that

$$I \vdash \mu' \Leftarrow x{:}A.\, S$$

case $\mu = (\chi, l \mapsto_\tau M), (z \Leftarrow !_\tau l; E_1), y{:}B.\, E_2$. By the typing of μ, we know

$$[\![\chi, l \mapsto_\tau M]\!] \vdash z \Leftarrow !_\tau l; E_1 \Leftarrow y{:}B.\, \nabla(I * R)$$

and

$$y{:}B; I; R \vdash E_2 \Leftarrow x{:}A.\, \nabla S.$$

By the typing of lookup

$$[\![\chi, l \mapsto_\tau M]\!] \wedge l \hookrightarrow_\tau z \vdash E_1 \Leftarrow \nabla(I * R)$$

After substituting z by M, we get

$$[\![\chi, l \mapsto_\tau M]\!] \wedge l \hookrightarrow_\tau M \vdash [M/z]E_1 \Leftarrow \nabla(I * R)$$

and then by consequence

$$[\![\chi, l \mapsto_\tau M]\!] \vdash [M/z]E_1 \Leftarrow \nabla(I * R)$$

Since $\mu' = (\chi, l \mapsto_\tau M), [M/z]E_1, y{:}B.\, E_2$, we can assemble back that

$$I \vdash \mu' \Leftarrow x{:}A.\, S.$$

case $\mu = \chi, \overline{\kappa_i}\ \overline{P_i}(\mathsf{publish}\ J; E)$. By typing of μ, we know that $\chi = \chi_1, \chi_2$ and $[\![\chi_1]\!] \vdash I$ and

$$I; [\![\chi_2]\!] \vdash \overline{\kappa_i}\ \overline{P_i}\ (\mathsf{publish}\ J; E) \Leftarrow x{:}A.\, \nabla S$$

From this judgment, we get

$$[\![\chi_2]\!] \vdash J * \top$$

and by inversion lemma

$$I * J; [\![\chi_2]\!] \circ (J \multimap \nabla\mathsf{emp}) \vdash \overline{\kappa_i}\ \overline{(P_i \circ (J \multimap \nabla\mathsf{emp}))}\ E \Rightarrow x{:}A.\, T'$$

where $y{:}B; \delta[\![\chi_2]\!] \circ (J \multimap \nabla\mathsf{emp}) \circ T' \vdash \nabla S$. Since ∇S does not depend on i, we can arbitrarily set the value of i on the left side of this judgment. In particular, we have:

$$\delta([\![\chi_2]\!] \circ (J \multimap \nabla\mathsf{emp})) \circ T' \vdash \nabla S$$

and therefore

$$I * J; [\![\chi_2]\!] \circ (J \multimap \nabla\mathsf{emp}) \vdash \overline{\kappa_i}\ \overline{(P_i \circ (J \multimap \nabla\mathsf{emp}))}\ E \Leftarrow x{:}A.\, \nabla S$$

From the fact $[\![\chi_2]\!] \vdash J * \top$ we know we can split χ_2 into $\chi_2 = \chi_3, \chi_4$ such that $[\![\chi_3]\!] \vdash J$. Furthermore, by preciseness of J, we know that χ_3 is the unique subheap of χ_2 satisfying J, and thus $[\![\chi_4]\!] \vdash [\![\chi_2]\!] \circ (J \multimap \nabla\mathsf{emp})$. Therefore, $\chi = \chi_1, \chi_3, \chi_4$ and $[\![\chi_1, \chi_3]\!] \vdash I * J$ and

$$I * J; [\![\chi_4]\!] \vdash \overline{\kappa_i}\ \overline{(P_i \circ (J \multimap \nabla\mathsf{emp}))}\ E \Leftarrow x{:}A.\, \nabla S$$

Since $\mu' = \chi, \overline{\kappa_i}\ \overline{(P_i \circ (J \multimap \nabla\mathsf{emp}))}\ E$, this suffices to assemble $I \vdash \mu' \Leftarrow x{:}A.\, S$.

case $\mu = \chi, \mathsf{return}\ M, y{:}B.\, E_2$. By typing of μ, we know

$$\chi \vdash \mathsf{return}\ M \Leftarrow y{:}B.\, \nabla(I * R)$$

and

$$y{:}B; I; R \vdash E_2 \Leftarrow x{:}A.\, \nabla S$$

By typing of return, $M \Leftarrow B$ and $[\![\chi]\!] \vdash I * R$. Thus, we can split χ into $\chi = \chi_1, \chi_2$ such that $[\![\chi_1]\!] \vdash I$ and $[\![\chi_2]\!] \vdash R$. After substituting y by M in the above typing judgment, and by the consequence principle, we get

$$I; [\![\chi_2]\!] \vdash [M/y]E_2 \Leftarrow x{:}A.\, \nabla S$$

Since $\mu' = \chi, [M/y]E_2$, this suffices to assemble $I \vdash \mu' \Leftarrow x{:}A.\, S$, as required.

■

Theorem 13 (Progress) *If $I \vdash \mu \Leftarrow x{:}A.\, S$, then either $\mu = (\chi, \mathsf{return}\ M)$, or there exists μ' such that $\mu \hookrightarrow \mu'$.*

Proof: The proof is by straightforward case analysis on μ.

The cases when μ is in a concurrent configuration do not pose any problems, basically because they never impose any changes on the heap, so there is always a μ' to be found. The cases when μ is in the atomic mode are also not hard, and are similar to the proofs presented in our previous work [20]. These cases rely on the following property of the assertion logic, which in the previous work we called "Heap Soundness".

1. If $[\![\chi]\!] \vdash l \mapsto_\tau -$ then $\chi = \chi_1, l \mapsto_\tau M$ for some χ_1, M.
2. If $[\![\chi]\!] \vdash l \mapsto -$ then $\chi = \chi_1, l \mapsto_\tau M$ for some χ_1, τ, M.

Heap soundness follows from the soundness of the assertion logic, which, in turn, is proved by building a simple denotational model of the calculus of constructions extended with monads, as we have done in Theorem 2. ■

7 Related and future work

Transactional Memory. Monads for dealing with transactions have been introduced in Haskell [8, 9]. Our approach is similar, however, we have not considered an explicit abort in this paper because we are interested in a high-level semantics where an explicit abort is not necessary [18]. Also, we can state and check the pre-conditions for an atomic block statically, and do not require an explicit abort to ensure correctness of algorithms.

Higher-order and dependent types for effects. Dependently typed systems with stateful features have to date mostly focused on how to appropriately restrict effects from appearing in types, thus polluting the underlying logical reasoning. Such systems have mostly employed singleton types to establish the connection between the pure and the impure levels of the language. Examples include Dependent ML by Xi and Pfenning [31, 32], Applied type systems by Chen and Xi [4] and Zhu and Xi [33], a type system for certified binaries by Shao et al. [27], and the theory of refinements by Mandelbaum et at. [13]. HTT differs from these approaches, because we allow effectful computations to freely appear in types, as the monadic encapsulation facilitates hygienic mixing of types and effects, and thus preserves soundness. There are also several recent proposals for purely functional languages with dependent types, like Cayenne [1], Epigram [15], Omega [28] and Sage [7]. We also list several works that extend Hoare and Separation Logics with higher-order functions, like the work of Honda, Berger and Yoshida [2] and Krishnaswami et al. [11]. To our knowledge, none of the mentioned languages and logics has been extended to a concurrent setting.

Separation Logic and concurrency. Resource invariants in (sequential) Separation Logic were introduced by O'Hearn et al. [23], and an extension to concurrency with shared resources has been considered by Brookes [3] and O'Hearn [22]. These works point out the need for precise invariants on the shared resources, in order to preserve the soundness of the logic. More recently, Vafeiadis and Parkinson [29] and Feng at el. [6] have combined Separation Logic with rely-guarantee reasoning, whereby processes specify upper and lower bounds on the evolution of the shared state.

Our treatment of shared state with invariants was inspired by O'Hearn's presentation in [22]. Using invariants simplifies the reasoning, but seems strictly weaker than rely-guarantee. Invariants only enforce a predetermined property, but otherwise lose information about the actual changes to the shared state. We have found this property somewhat restrictive in several examples, and plan in the future to reformulate HTT with ideas from the rely-guarantee approaches.

Implementation and models of concurrency. The model of HTT described here suffices to argue soundness, but is otherwise quite restrictive, as it cannot support any interesting relations on effectful computations, except the monadic laws. A more refined model of sequential, impredicative, HTT has been developed by Petersen et al. [25]. We hope that this model can be extended to a setting with transactions as well.

To improve usability of HTT, we plan to support automatic inference of (some) pre- and post-conditions and loop invariants. This would avoid the current need to explicitly annotate the concurrency primitives. Currently, HTT rules compute strongest postconditions, but a significant amount of annotations can be inferred if the rules are re-formulated to simultaneously infer the weakest precondition that guarantees progress, as well as the strongest postcondition with respect to this precondition. We are currently implementing this kind of formulation as an extension of Coq, supporting evaluation as code extraction into Haskell.

8 Conclusion

This paper presented Hoare Type Theory (HTT), which is a dependently typed programming language and logic supporting higher-order programs with transactional shared memory concurrency.

HTT follows the "specifications-as-types" principle, and internalizes specifications in the form of Hoare triples for partial correctness of stateful and concurrent programs into types. This isolates the concerns

about side-effects and concurrency from the logical, purely functional foundations of the system, and makes it possible to mix concurrency with various advanced features, like higher-order functions, polymorphism, ADTs, none of which was possible in the previous work on Hoare or Separation Logics for concurrency. In fact, the pure fragment of HTT can soundly be scaled to the Extended Calculus of Constructions ECC [12] and Coq [14].

Hoare specifications in HTT are monads, and we support two different monadic families: $\mathsf{ST}\,\{P\}\,x{:}A\{Q\}$ classifies stateful sequential computations where P and Q are pre- and post-conditions on the state, and $\mathsf{CMD}\,\{I\}\{P\}\,x{:}A\{Q\}$ classifies transactional computations, where I is an invariant on the shared state and P and Q are pre- and post-condition on the local state. Both monads use propositions from Separation Logic to concisely describe the various aspects of the process state. Transactional computations may *atomically* invoke a stateful computation on the *shared state*, if the stateful computation provably preserves the invariant of the shared state. That is, we provide a primitive atomic, which can coerce the type $\mathsf{ST}\,\{P * I\}\,x{:}A\{Q * I\}$ into the type $\mathsf{CMD}\,\{I\}\{P\}\,x{:}A\{Q\}$.

We have shown that HTT as a logic is sound and compositional, so that it facilitates local reasoning. We have defined its operational semantics, and shown that this semantics is adequate with respect to the specifications from the Hoare types.

References

[1] L. Augustsson. Cayenne – a language with dependent types. In *International Conference on Functional Programming, ICFP'98*, pages 239–250, 1998.

[2] M. Berger, K. Honda, and N. Yoshida. A logical analysis of aliasing in imperative higher-order functions. In O. Danvy and B. C. Pierce, editors, *International Conference on Functional Programming, ICFP'05*, pages 280–293, Tallinn, Estonia, September 2005.

[3] S. Brookes. A semantics for concurrent separation logic. In *CONCUR'04*, volume 3170 of *Lecture Notes in Computer Science*, pages 16–34, 2004.

[4] C. Chen and H. Xi. Combining programming with theorem proving. In *International Conference on Functional Programming, ICFP'05*, pages 66–77, Talinn, Estonia, September 2005.

[5] E. W. Dijkstra. Guarded commands, nondeterminacy and formal derivation of programs. *Communications of the ACM*, 18(8):453–457, 1975.

[6] X. Feng, R. Ferreira, and Z. Shao. On the relationship between concurrent separation logic and assume-guarantee reasoning. In *European Symposium on Programming, ESOP'07*, volume 4421 of *Lecture Notes in Computer Science*, pages 173–188, 2007.

[7] C. Flanagan. Hybrid type checking. In *Symposium on Principles of Programming Languages, POPL'06*, pages 245–256, 2006.

[8] T. Harris and K. Fraser. Language support for lightweight transactions. *ACM SIGPLAN Notices*, 38(11):388–402, Nov. 2003.

[9] T. Harris, S. Marlow, S. L. P. Jones, and M. Herlihy. Composable memory transactions. In *Symposium on Principles and Practice of Parallel Programming, PPoPP'05*, pages 48–60, 2005.

[10] M. Hofmann. *Extensional Concepts in Intensional Type Theory*. PhD thesis, Department of Computer Science, University of Edinburgh, July 1995. Avaliable as Technical Report ECS-LFCS-95-327.

[11] N. Krishnaswami. Separation logic for a higher-order typed language. In *Workshop on Semantics, Program Analysis and Computing Environments for Memory Management, SPACE'06*, pages 73–82, 2006.

[12] Z. Luo. *An Extended Calculus of Constructions*. PhD thesis, University of Edinburgh, 1990.

[13] Y. Mandelbaum, D. Walker, and R. Harper. An effective theory of type refinements. In *International Conference on Functional Programming, ICFP'03*, pages 213–226, Uppsala, Sweden, September 2003.

[14] The Coq development team. *The Coq proof assistant reference manual*. LogiCal Project, 2004. Version 8.0.

[15] C. McBride and J. McKinna. The view from the left. *Journal of Functional Programming*, 14(1):69–111, January 2005.

[16] J. L. McCarthy. Towards a mathematical science of computation. In *IFIP Congress*, pages 21–28, 1962.

[17] E. Moggi. Notions of computation and monads. *Information and Computation*, 93(1):55–92, 1991.

[18] K. F. Moore and D. Grossman. High level small step operational semantics for transactions. In *Workshop on Transactional Computing*, August 2007.

[19] A. Nanevski, A. Ahmed, G. Morrisett, and L. Birkedal. Abstract Predicates and Mutable ADTs in Hoare Type Theory. In *European Symposium on Programming, ESOP'07*, volume 4421 of *Lecture Notes in Computer Science*, pages 189–204, 2007.

[20] A. Nanevski, G. Morrisett, and L. Birkedal. Hoare Type Theory, Polymorphism and Separation. *Journal of Functional Programming*, pages ??–??, 2007. To appear.

[21] P. O'Hearn, J. Reynolds, and H. Yang. Local reasoning about programs that alter data structures. In *International Workshop on Computer Science Logic, CSL'01*, volume 2142 of *Lecture Notes in Computer Science*, pages 1–19. Springer, 2001.

[22] P. W. O'Hearn. Resources, concurrency and local reasoning. *Theoretical Computer Science*, 375(1–3):271–307, May 2007.

[23] P. W. O'Hearn, H. Yang, and J. C. Reynolds. Separation and information hiding. In *Symposium on Principles of Programming Languages, POPL'04*, pages 268–280, 2004.

[24] S. Owicki and D. Gries. Verifying properties of parallel programs: an axiomatic approach. *Communications of the ACM*, 19(5):279–285, 1976.

[25] R. L. Petersen, L. Birkedal, A. Nanevski, and G. Morrisett. A realizability model for impredicative Hoare Type Theory. Unpublished manuscript, July 2007.

[26] F. Pfenning and R. Davies. A judgmental reconstruction of modal logic. *Mathematical Structures in Computer Science*, 11(4):511–540, 2001.

[27] Z. Shao, V. Trifonov, B. Saha, and N. Papaspyrou. A type system for certified binaries. *ACM Transactions on Programming Languages and Systems*, 27(1):1–45, January 2005.

[28] T. Sheard. Languages of the future. In *Object-Oriented Programming, Systems, Languages, and Applications, OOPSLA'04*, pages 116–119, October 2004.

[29] V. Vafeiadis and M. Parkinson. A marriage of rely/guarantee and separation logic. In *CONCUR'07*, 2007. to appear.

[30] K. Watkins, I. Cervesato, F. Pfenning, and D. Walker. A concurrent logical framework: The propositional fragment. In S. Berardi, M. Coppo, and F. Damiani, editors, *Types for Proofs and Programs*, volume 3085 of *Lecture Notes in Computer Science*, pages 355–377. Springer, 2004.

[31] H. Xi and F. Pfenning. Eliminating array bound checking through dependent types. In *Conference on Programming Language Design and Implementation, PLDI'98*, pages 249–257, Montreal, Canada, June 1998.

[32] H. Xi and F. Pfenning. Dependent types in practical programming. In *Symposium on Principles of Programming Languages, POPL'99*, pages 214–227, January 1999.

[33] D. Zhu and H. Xi. Safe programming with pointers through stateful views. In *Practical Aspects of Declarative Languages, PADL'05*, volume 3350 of *Lecture Notes in Computer Science*, pages 83–97, Long Beach, California, January 2005. Springer.