Types for Relaxed Memory Models

Matthew Goto  Radha Jagadeesan  Corin Pitcher  James Riely
DePaul University

Abstract
Multicore computers implementing weak memory models are mainstream, yet type-based analyses of these models remain rare. We help fill this gap. We not only prove the soundness of a type system for a weak execution model, but we also show that interesting properties of that model can be embedded in the types themselves.

We argue that correspondence assertions can be used in a programming discipline that captures happens-before relationships, which are the basis for reasoning about weak memory in Java. This programming discipline is flexible and can be statically enforced. We present several examples from java.util.concurrent and prove the static semantics sound with respect to an execution model based on Java’s memory model.

Categories and Subject Descriptors D.3.1 [Programming Languages]: Formal Definitions and Theory

General Terms Theory, Languages

1. Introduction

The C++ memory model posits that programs with data races have undefined semantics [Boehm and Adve 2008]. This model is incompatible with languages such as Java, which allow data races but seek to ensure safety. The Java Memory Model (JMM) [Manson, Pugh, and Adve 2005] is a speculative model intended to validate the executions generated by modern processors and compilers while banning thin-air reads, and thus to ensuring type safety. To our knowledge, the safety argument has only been made informally.

The happens-before relation is crucial to understanding shared-memory concurrent Java programs. The Java APIs contain extensive comments describing the happens-before relation, such as the following: “Actions in a thread prior to placing an object into any concurrent collection happen-before actions subsequent to the access or removal of that element from the collection in another thread.” There is no method for describing happens-before relations formally.

Consider the following tiny example, which shows that, without proper synchronization, it is possible for a thread to read stale values. We adopt the following notational conventions: \( a \) and \( t \) are thread identifiers; \( p \) is an object identifier; the process \( a[M] \) denotes a thread with identifier \( a \), executing statement \( M \); the process \( AB \) denotes the parallel composition of threads \( A \) and \( B \).

```java
class C1 { int f; int g; }
s[p.f=42; p.g=27;]  |
t[val y=p.g; val x=p.f; if((y!=0)&&(x==0)){†}]
```

There are two threads sharing object \( p \) inhabiting class \( C1 \), (notation \( p:C1 \)). All fields are initially set to 0. Object \( p \) has two fields. Thread \( s \) updates \( p.f \) then \( p.g \). Thread \( t \) reads the fields in the reverse order. Because there is no synchronization, the program has data races on both fields.

In a sequentially consistent execution [Lamport 1979], threads must appear to execute sequentially in an interleaved fashion. Since \( s \) must write \( p.f \) before \( p.g \), it is impossible for \( t \) to read an object from \( p.g \) and subsequently read \( 0 \) from \( p.f \); thus the statement \( † \) is unreachable in a sequentially consistent semantics.

However, compilers may re-order unrelated writes and reads, and thus the JMM semantics is designed to allow executions in which \( † \) is reachable in example \( C1 \).

Synchronization eliminates data races, shrinking the set of visible values. For example, in the following variant of example \( C1 \), field \( g \) is declared as volatile. There are happens-before edges from a write of a volatile to every subsequent read. Assume \( p:C2 \).

```java
class C2 { int f; volatile int g; }
s[p.f=42; p.g=27;]  |
t[val y=p.g; val x=p.f; if((y!=0)&&(x==0)){†}]
```

Because of the synchronization though \( p.g \), \( † \) becomes unreachable in the JMM for \( C2 \). Such happens-before reasoning is subtle: in the variant where \( f \) is volatile but \( g \) is not, \( † \) is reachable. The reasoning may also fail in the presence of other threads.

In this paper we propose a a compositional program discipline for happens-before using correspondence assertions [Woo and Lam 1992; Gordon and Jeffrey 2003]. As a byproduct, we present a proof of type safety for concurrent Java.

Organization. In Section 2, we sketch the semantics of our memory model and present correspondence assertions as a tool for reasoning about happens-before relations. Section 3 is a brief detour to complete the description of the semantics; we introduce speculation and describe its implications for type soundness. In Section 4, we provide intuitions about the type system via examples. For readability, Sections 1–4 sometimes use Java syntax. For simplicity of formal development, the remainder of the paper uses a more restrictive language.

In Section 5, we define the static type system, which is used to validate class definitions. In Section 6, we define soundness using an operational semantics. In Section 7, we describe the soundness proof, highlighting the most interesting cases of the inductive invariant maintained by the operational semantics. In Section 8, we summarize related work and conclude.

Details of the soundness proof, including the typing rules for processes used in subject reduction, are available in the full version of the paper.
2. Happens-before semantics using correspondence assertions

We provide a rough introduction to our operational semantics and describe how correspondence assertions can be integrated to reason about happens-before relations. The semantics is essentially copied from our previous paper [Jagadeesan, Pitcher, and Rieyl 2010], although it has been adapted to handle atomics.

The semantics is based on the idea that writes do not affect a traditional store, but rather produce actions. Although it has been adapted to handle atomics, the form of an atomic denotation is $\text{Atomic}(v; j)$, where $v$ is the current value and $j$ is the opcount.

We discuss synchronization order using Java’s atomic objects, rather than volatiles. A volatile field can be viewed as a final atomic upon which we only call get and set. In addition, atomics have get-and-set and compare-and-swap methods, which can be used to implement locks.

Synchronization order specifies that a “set” on an atomic happens-before any subsequent “get.” We model this using synchronization actions of the form $(\ell s t j)$, where $\ell$ the atomic object acted upon, $s$ is the acting thread, and $j$ is an opcount. If the opcount is odd, then the operation was a set. If the opcount is even, then the operation was a get. The opcount is incremented when a set follows a get or a get follows a set. Initially, atomics are “set” with opcount 1.

Reading an atomic is deterministic; thus we store the value held by an atomic with its denotation, rather than representing it using actions. The form of an atomic denotation is $\text{Atomic}(v; j)$, where $v$ is the current value and $j$ is the opcount.

Consider the following variant of example C2 from the introduction.

$p: \text{Int} \mid l: \text{Atomic}(0; 1) \mid s[p.f=42; l.set(27)] \mid t[\text{val y=l.get(); \text{val x=p.f; if((x==0)&&(y!=0)){};}]}$

Just as for example C1, statement $t$ is unreachable. For example, consider the case where thread $a$ executes before $t$. In this case, the read of $p.f$ occurs in the following context.

$(s!p.f=42)!s[1:1] / (t!1:2)
(p: \text{Int} \mid l: \text{Atomic}(27; 2) \mid s[\cdot] \mid [-])$

The synchronization action $(s[1:1])$ was created when a performed $1.set(27)$; this operation modified the value of the atomic, but not its opcount, resulting in $l: \text{Atomic}(27; 1)$. The next synchronization action, $(t!1:2)$, was created when $t$ performed $1.get()$; this operation modified the opcount of the atomic, but not its value, resulting in $l: \text{Atomic}(27; 2)$. Note that synchronization actions record the opcount, whereas write actions, such as $(s!p.f=42)$, record the value written.

The above context has three happens-before edges, when interpreted as a context for $t$.

* program order from $(s!p.f=42) \rightarrow (s[1:1])$
* synchronization order from $(s[1:1]) \rightarrow (t!1:2)$, and
* program order from $(t!1:2) \rightarrow$ the hole.

Happens-before is also affected by thread creation. For example, the bootstrap code given above will evaluate to the following context, holding process $s[\cdot\cdot\cdot] \mid t[\cdot\cdot\cdot]$.

$(\text{main}) / (\text{main})
(m!p.f=0) / (m!1:1) / (m!k:1) / (s[1:2]) / (t!k:2)
(m: \text{Main} \mid p: \text{Int} \mid s: \text{S} \mid t: \text{T} \mid [-])$

Given the semantics of $\text{start}$, there are synchronization edges from $n$ to the threads that it creates, here manifest in the synchronization variables $l$ and $k$, which will never be used again. Because of these synchronization edges, the write of the initial value of $p.f$ happens-before any actions performed by $s$ or $t$, since these threads are part of the process filling the hole in the context.

Many concurrent algorithms use locks. When locks are implemented using atomics, the release of a lock happens-before the subsequent acquire.

In this paper we reason about happens-before relations using correspondence assertions [Woo and Lam 1992; Gordon and Jeffrey 2003]. Using this technique, a programmer may annotate a program with begin and end statements. The execution of an end statement is correct if it occurs after the corresponding begin. The primary goal of static analysis is to ensure that all ends are correct.

Acknowledgements. We gratefully acknowledge the detailed comments of the anonymous referees on this and a previous version of the paper. This research was supported by NSF grant #0916741.
In a well-typed program, therefore, end statements can be erased without changing the meaning of the program.

Correspondences have no computational effect. The particular form of correspondence is chosen in order to specify some property of interest. We wish to reason about happens-before relations in code built using atomics and other mechanisms, and we wish to do so in a way that might be useful for API documentation.

Recall the example from java.util.concurrent quoted in the introduction: "Actions in a thread prior to placing an object into any concurrent collection happen-before actions subsequent to the access or removal of that element from the collection in another thread." Directly capturing such statements requires dependent types, as well as a selection of analyses based on locality [Sewell 1998], confinement [Zhao, Palsberg, and Vitek 2006], linearity [Hawblitzel 2005], and effects [Lucassen and Gifford 1988]. Such powerful type systems provide strong guarantees at the cost of additional programmer effort that we do not wish to mandate.

Our goal has been to develop a type system that captures some property of happens-before while retaining simple types. To achieve this, we propose a rather weak notion of correspondence.

In our minimalist approach, the creation of an object \( p \) by a thread \( s \) begins a correspondence, creating a begin action (\( s!p \)). At any point, a program can attempt to end the correspondence using the statement \( \text{end } p \). If (\( s!p \)) happens-before the hole in the context where \( p \) occurs, then we say that this occurrence of \( p \) is correctly published, and the end acts as a no-op; otherwise the thread becomes stuck. For example, consider the following process.

\[
\langle s!p \rangle \langle s[\text{end } p; \cdots] \text{ } t[\text{end } p; \cdots] \rangle
\]

Thread \( s \) evaluates to \( s[\cdots] \) whereas thread \( t \) is stuck: the begin in context (\( s!p \)) [-] only happens-before the hole when the hole is filled by \( s \). A stuck program is erroneous; typing ensures that well-typed programs cannot get stuck.

This form of specification is weak in the sense that a programmer cannot choose the point at which a begin occurs; begins always coincide with object creation. The programmer is also restricted to end on a single object reference, not a tuple.

Additionally, Java allows the null value wherever an object reference occurs, not a tuple. For example, consider the following process.

\[
\text{atomic} \to (s!p.f=42)(s[\text{end } o;] \text{ } t[\text{val } x=l.get(); \text{end } x;])
\]

This case, the residual \( t[\text{end } o;] \) is executed in the context \( (s!o) (s!p.f=27) \) and therefore is not stuck.

Example C1 from the introduction can be rewritten to use correspondences as follows, where p:C1'

\[
\text{class C1' \{ Object f; Object g; \}}
\]

\[
\text{s!p.f=new Object(); p.g=new Object() \}} \text{ \text{ | t[\text{val } y=p.g; \text{val } x=p.f; \}}
\]

\[
\text{if(y!=null){\text{end } y; \text{if(x==null){}}}]
\]

In C1', the end assertion can get stuck, since there is no happens-before edge from \( s \) to \( t \).

Example C2 can also be written in this form, where p:C2'

\[
\text{class C2' \{ Object f; final Atomic g; \}}
\]

\[
\text{s!p.f=new Object(); p.g.set(new Object()) \}} \text{ \text{ | t[\text{val } y=p.g.get(); \text{val } x=p.f; \}}
\]

\[
\text{if(y!=null){\text{end } y; \text{if(x==null){}}}]
\]

In C2', the end assertion cannot get stuck, and statement \( \dagger \) is unreachable. If the reads of \( t \) occur in the reverse order, then \( t \) could become stuck and statement \( \dagger \) could be unreachable.

### 3. Speculation

A substantial technical hurdle in proving soundness of any type system for the JMM is the treatment of speculation. We first describe our semantics for speculation, then discuss the implications for the proof. See our previous paper [Jagadeesan, Pitcher, and Riely 2010] for details.

While the JMM is clearly designed with type safety in mind, it does not lend itself to standard proofs using subject reduction. The JMM validates an execution using a sequence of commitments and a corresponding sequence of validating executions: a commitment \( C \) is a set of actions resolving some read-write data races; in a sequence of valid executions, execution \( E \) must justify \( C \) using \( C_{n-1} \), subject to various coherence criteria. The coherence criteria for the JMM are notoriously subtle, and various revisions have been proposed [Cenciarelli, Knapp, and Sibilio 2007; Sevcik and Aspinall 2008; Sevcik 2008]. These coherence criteria complicate the normal inductive argument for soundness, perhaps accounting for the fact that no soundness proof has been published as yet (to the best of our knowledge).

For our soundness proof, we adopt the semantics proposed in our previous work, which handles speculation quite differently. Our semantics allows every execution allowed by the JMM, with the exception of some programs that contain both synchronization actions and data races. As a result of these subtle differences, our semantics validates many common program transformations that the JMM was meant to validate, but does not.

In addition to the mechanisms present in the previous section for fields and atomics, the semantics allows single threaded action rewriting and speculation.

Single threaded action rewriting is easy to understand. We allow dynamic transformations to the action sequence generated by a single thread, as long as this does not introduce new behaviors. For example, it is permitted to rewrite (\( s!p.f=42 \))\( \langle s!p.f=27 \rangle \) to \( \langle s!p.f=27 \rangle \), removing the value 42, which may be visible to concurrent threads. The converse transformation is not sound, however, since it introduces the value 42 out of thin-air.

Atomics do preserve safety. Consider the following program, derived from the previous one by replacing \( p.f \) with \( \text{atomic } 1 \).

\[
\langle s!o \rangle (s!1:1)(1:\text{Atomic}(o);1) \text{ } s[\text{end } o;] \text{ } t[\text{val } x.l.get(); \text{end } x;])
\]

In this case, the residual \( t[\text{end } o;] \) is executed in the context \( (s!o) (s!1:1) t[1:2] \) and therefore is not stuck.

Example C1 from the introduction can be rewritten to use correspondences as follows, where p:C1'

\[
\text{class C1' \{ Object f; Object g; \}}
\]

\[
\text{s!p.f=new Object(); p.g=new Object();} \text{ \text{ | t[\text{val } y=p.g; \text{val } x=p.f;}\]
\]

\[
\text{if(y!=null){\text{end } y; \text{if(x==null){}}}]
\]
Speculation is far more subtle. Let "\(\exists x\)" stand for return \(x\) in the following example.

\[
\{\text{val } x = p \cdot g; \text{ p.f=27}; \exists x \} \quad \text{t} \{\text{val } y = p \cdot f; \text{ p.g=y}\}
\]

In any sequentially consistent execution, \(s\) must read the initial value 0 from \(p \cdot g\) resulting in \(s\[0\]. A speculative semantics also allows the result \(s\[27\]. In an implementation, such an execution may be the result of cache effects or optimizations that reorder the independent statements of \(s\).

The basic idea of speculation is to allow \(s\) to read the value 27 from \(p \cdot g\) before it is written by \(t\). Clearly, we are playing with knives. It is common to impose two sanity checks on speculation. First, it should not create data out of “thin air”. Second, any program without data races should only have sequentially consistent executions. Our semantics guarantees both properties by restricting valid speculations. We refer to our previous paper for full details, providing only a flavor of the ideas here.

Speculation is not performed actively by a process. Rather, a speculation happens to a process. Speculation is performed, non-deterministically, by the reduction relation itself. For example, let \(A\) be the original pair of threads given above. We have the following “reduction,” which speculates that \(t\) will write 27 to \(p \cdot g\).

\[
A \rightarrow \left( \left( T \Rightarrow A \right) \Leftarrow \left( \left( t \cdot p \cdot g=27 \Rightarrow A \right) \right) \right)
\]

Speculative “actions” are marked with a query (?), whereas concrete action are marked with a bang (!). The reduction creates two copies of the original process, which are executed in separate universes with separate copies of the state. The left copy is called the initial process; the right, the final process.

The initial and final processes evolve independently. Within the final process, any thread but \(t\) can read the speculative value. This allows reduction to the following.

\[
T \Rightarrow \left( \left( t ! p \cdot g=27 \right) \left( t ! p \cdot g=27 \right) \left( s[\exists 0] \parallel t [] \right) \right)
\]

Actions performed by both processes are allowed to trickle out of the speculation, in order. In addition, object denotations and threads can move in and out of the speculation, as long as there are identical copies on both sides. In the case of our example, this results in

\[
\left( s ! p \cdot f=27 \right) \left( t ! p \cdot g=27 \right) \left( s[\exists 27] \parallel t [] \right)
\]

When a speculation occurs in a context with a matching concrete write, the speculation may be removed, leaving only the final process; the initial process is used only to justify the action that removes the speculation. This is called finalization. For our running example, the result is

\[
\left( s ! p \cdot f=27 \right) \left( t ! p \cdot g=27 \right) \left( s[\exists 27] \parallel t [] \right)
\]

A valid execution is one in which every speculation can be finalized; non-valid executions may create values out of thin air, and, therefore, are not interesting. Nondeterminism ensures that the reduction relation captures all valid executions.

The soundness theorem establishes that typed programs cannot become stuck and therefore that every end \(p\) occurs in a context such that, for some \(s\), \(s ! t p\) happens-before the context hole.

Soundness only applies to valid executions. The difficulty is to find an invariant that holds in the face of partial executions which have no valid extension. Such partial executions are caused by speculations which cannot be finalized, which we refer to as traps.

If a trap is the result of shape error, our type system can safely treat the final branch as untyped: we simply “turn off” typing in worlds that we know can never be exposed at top level.

Our type system includes effects, which formalize happens-before relations using correspondence assertions. Thus, in addition to shape errors, there are also subtle traps, such as speculations that occur without a thread holding appropriate locks. Our proof addresses this subtlety by being safely optimistic about effects: if the optimism is justified, then it has done no harm; if the optimism is not justified, then the speculation must be a trap. In either case, subject reduction holds. For further discussion, see Section 7.1.

It may be argued that we should simply stop the reduction relation from generating unfinalizable speculations. In general, however, unfinalizability is undecidable. Of course, we can rule out certain classes of speculation which are clearly unfinalizable, including shape errors and locking errors. However, we believe that in order to know that a speculation is unfinalizable, you must allow the speculation to occur, and then prove that it cannot be finalized.

4. Static approximation of happens-before

Recall from Section 2 that an occurrence of an object reference is correctly published if an end statement using that occurrence cannot become stuck. An end statement that cannot become stuck in any execution is safe. We introduce typing annotations to statically approximate correct publication.

Correct publication is a predicate on occurrences of object references, not on objects themselves: threads may have different happens-before relations to the begin action that marks the creation of an object. A copy of a correctly published reference need not itself be correctly published. In particular, correct publication may be lost when a reference is copied via a shared object. Communication using an atomic preserves correct publication, since an atomic read is related by happens-before to every write to the same atomic. However, communication using a non-atomic does not necessarily preserve correct publication.

A field or variable annotated as reliable (notation \(\mathbb{R}\)) must hold only correctly published references. Annotations may be attached to any field or variable, including method formal parameters, the implicit parameter this, and return values. The key to safely approximating correct publication in the type system is to ban read-only data races on reliable fields. We consider three mechanisms for doing so: final fields, atomics and lock-protected fields. We discuss these three primitives in more detail below. In the following subsections, we discuss larger data structures using these primitives.

Final fields. Final fields have a special semantics in the JVM and can be used to communicate correct publication. If field \(f\) is final and reliable, then \(p.f\) is correctly published if the occurrence of \(p\) is correctly published.

Atomics. We take Atomic to be primitive in the dynamics, and provide it an enriched signature in the statics to indicate that it preserves happens-before relations. We first present a simplified version of the signature for Atomic.

```java
class Atomic {
    Object get();
    void set(Object v); ...
}
```

In the case of get, the reliable annotation indicates that a client program can safely end on the returned value. In the case of set, the annotation requires that all values written to an atomic be correctly published. This restriction may be undesirable in some cases, and therefore we introduce polymorphism over annotations. The final form of the atomic interface is as follows.

```java
class Atomic(CH) {
    CH Object get(CH this); // or CH Object get();
    void set(CH this; CH Object v); ...
}
```

Here \(CH\) is an annotation variable, which may be instantiated to \(\mathbb{R}\) or \(\square\), indicating whether a variable is reliable. Reliable variables must hold only correctly published values, whereas variables that are not reliable variables may hold any value.
It is always safe to forget that a value is correctly published, thus losing the capability to end. Annotations therefore give rise to a natural form of subtyping with $\mathbb{X} < \chi < \triangleright$. In general, annotations are required on the implicit parameter this. Here the annotations indicate that correctly published values may be recovered from an Atomic that is not itself correctly published, that is, this is marked $\triangleright$.

We elide $\square$ annotations except when they occur as actual parameters to generic classes. We also elide the this annotation in method declarations when it is $\square$. This example C2' from the end of Section 2 in our system, we must specify that the atomic held in g is reliable. The correct declaration of C2' is as follows.

```java
class C2' { Object f; final Atomic g; }
s[p.f=new Object(); p.g.set(new Object());] | t[val=y.p.g.get(); val=x.p.f; if(y!=null){end y; if(x=null){t}}]
```

In this example, there are two classes of executions: those in which t gets null from p.g and those in which t gets the object set by s. In the former case, the annotations on Atomic say nothing. In the latter case, the annotations indicate that the creation of the parameter to set happens-before-the return of get. That is, happens-before relations are preserved by Atomic from a set to the corresponding get.

Although we take Atomic to be a primitive, one can imagine varying implementations. Implementing Atomic by simply reading and writing a field would violate the interface given above, since such an implementation would allow a write before a set to race against a read after the corresponding get. However, an implementation that always returns null would satisfy the interface trivially. More subtly, so would an implementation that created a fresh object for each get, or always returned the first object passed to set. In both cases, the happens-before relation that results is not the expected one. In the former case, we have that the body of get happens-before the return of set. In the latter case, we have that the first call to set happens-before the return of every subsequent call to get. In short, the annotations say nothing about behavior of Atomic from a set to the corresponding get.

Using the data structure Atomic as a data structure. As a data structure, Atomic requires that after the first set, every get action returns the value from the "previous" set. Such requirements are readily understood informally, but difficult to formalize in a way that appeals to non-specialists.

Lock-protected fields. Communicating correct publication on fields which are neither final nor atomic is rather more delicate, as one must guarantee the absence of read-write data races programmatically. The use of locks to avoid races has been studied by Abadi, Flanagan, and Freund [2006]. We adapt the simplest of their systems to our setting by allowing write-write data races on all fields and read-write races on fields that are not reliable. Consider the following locking variant, where p:C3 and l:Lock. (Here synchronized statements do not create scopes, and therefore y remains visible.)

```java
class C3 { Object f; Object g guarded by 1; }
s[p.f=new Object();]
  synchronized(l){p.g=new Object();} | t[synchronized(l){val=y.p.g; val=x.p.f;
      if(y!=null){end y; if(x=null){t}}]
```

Just as in example C2' above, the end assertion is safe, and statement t is unreachable. In this case, the happens-before relation is carried on lock 1. The type of g indicates that it is a conservative field guarded by 1. A guarded field may only be accessed in a context that holds the corresponding lock, thus ensuring the necessary happens-before relations from lock acquisition to field read and from field write and lock release.

In general, we allow reads and writes to be guarded by different boolean combinations of locks, requiring that any set of locks satisfying the read guard must overlap with every set of locks satisfying the write guard. (Overlap is defined in Section 5.2 to capture the intuition of non-disjoint intersection.) This disallows read-write data races while still allowing write-write data races. Consider an example from Sevcik [2008] with write-write data races, but no read-write data races, where p:D:1:Lock, and k:Lock are shared.

```java
class D { Object f; guarded (1 && k) wrguard (1 || k); }
s[synchronized (1){p.f=new Object();}] | t[synchronized (k){p.f=new Object();}] | u[synchronized (1){
  synchronized (k){oObject x=p.f;}}]
```

The annotations indicate that the read of p.f requires both locks 1 and k, whereas the write of p.f requires only one of them. This program typechecks in our system. While this example is contrived, the separate treatment of read and write capabilities on objects is critical to the design of efficient data structures to avoid unnecessary sequentialization (see Java’s concurrent hash tables).

4.1 Dummy variables

Our technique uses object identity to reason about dynamic occurrences of program points via correspondence assertions. Occasionally, however, interfaces lack appropriate variables to annotate the happens-before relations of interest. In such cases, we require that the interface be augmented with dummy variables, which can be systematically erased at runtime. For example, consider the following methods of Atomic, with candidate annotations. (Recall that we elide the annotation for the implicit this from method types when annotation is $\square$.)

```java
class Atomic($\chi$) { ...
  $\chi$ Object getAndSet($\chi$ Object v);
  boolean compareAndSet(Object oldv, $\chi$ Object newv); }
```

Using the data structure invariants, the annotation of the getAndSet method captures exactly the right property: a return from getAndSet happens-after the previous set, and a call to getAndSet happens-before the subsequent get. The compareAndSet method has the same memory effect as getAndSet, however, its annotation only gives half of the story: the return value is unannotated. The situation can be remedied using a dummy variable in the signature of compareAndSet.

```java
Pair{boolean, $\chi$ Object}
  compareAndSet(Object oldv, $\chi$ Object newv);
```

If compareAndSet returns false, the dummy variable returns null; otherwise, the dummy variable returns the previous value stored by the atomic, as in getAndSet.

The unannotated interface of BinaryLatch (adapting Java’s CountDownLatch) is:

```java
class BinaryLatch { void await(); void countDown(); }
```
The documented memory effect is that a return of `await` happens-after the corresponding call to `countDown`. Applying our methodology leads to the interface

```java
class BinaryLatch<χ> { χ Object await();
  void countDown(χ Object v); }
```

where `await` returns the value passed to the last call to `countDown`. Similar changes can be made systematically throughout the APIs.

class Condition<χ> { χ Object await();
  void signal(χ Object); }

class Semaphore<χ> { χ Object acquire();
  void release(χ Object v); }

The annotations on `Condition` indicate the memory effect that a return of `await` happens-after the corresponding call to `signal`, and similarly for the (binary) `Semaphore`.

### 4.2 Building data structures

From the primitive constructs with special memory effects—final fields, atomic fields, and lock-protected fields—it is possible to construct useful concurrent data structures. The happens-before relations created by these data structures are key to concurrent programming. Indeed, every data structure provided by the Java `util.concurrent` API specifies such memory consistency effects.

Our typing annotations are a lightweight mechanism for making these memory effects amenable to the static analysis.

For example consider `BlockingQueue` from Java `util.concurrent`. The Java APIs specify the behavior of a `BlockingQueue` in two parts.

- The data structure specification states that there is a one-to-one correspondence between invocations of the `take` and the `put` methods, such that a `take` returns the `Object` that is added into the data structure by its `put` method.
- The memory consistency specification states that “actions in a thread prior to placing an object into a `BlockingQueue` happen-before actions subsequent to the access or removal of that element from the `BlockingQueue` in another thread”.

In our system, the `BlockingQueue` is annotated as follows.

```java
class BlockingQueue<χ> { void put(χ Object o);
  χ Object take();... }
```

As with `Atomic`, the annotations do not capture the data structure invariant, and only very roughly approximate the memory consistency specification. Nevertheless, they provide valuable information. Consider the following variant of our running example, where `p:C4` and `q:BlockingQueue(κ)`. The call to `take` blocks until `put` begins execution. Assuming that `q` is initially empty and that there are no other threads, then `t` must take the object that is put by `s`. In this case, the annotation tells us that the call to `take` happens-before the return of `get`. By thread-local reasoning, we can conclude that the writes by `s` must happen-before the call to `put`, and also that the reads by `t` must happen-after the return of `take`. Reasoning by transitivity, we can conclude that statement `†` is unreachable.

Formally, our system allows the inclusion of “end `z`” in thread `t`, indicating that the call of `put` happened-before the return of `take`. To reason about the reachability of `†`, one must also reason about thread-local happens-before relations and the confinement `q` and `p`. Our system captures the crucial cross-thread reasoning required about happens-before, leaving the local reasoning to informal methods or modular formal methods.

In the rest of this section, we present both non-blocking and lock-based examples which typecheck in our system, providing typed implementations of the interfaces discussed above.

### 4.3 Non-blocking algorithms

The basic building block for writing non-blocking algorithms is the `Atomic` class. For example, one can implement the `BinaryLatch` interface using an `Atomic`.

```java
class BinaryLatch<χ> {
  final Object init;
  final Atomic<χ> witness = new Atomic();
  BinaryLatch() {
    witness.set(new Object());
    init = witness.get();
  }

  χ Object await() {
    while (witness.get() == init);
    return witness.get();
  }

  void countDown(χ Object o) {
    witness.compareAndSet(init, o);
  }
}
```

Informally, one can reason that the return of `await` must happen-after the call to `compareAndSet` which must happen-after the corresponding call to `countDown`. Thus, the implementation validates the interface of `BinaryLatch`.

This guarantee is established formally by typechecking the code in our system. The type system uses the `Atomic` interface to deduce that the return value of `await` may be annotated by `χ`. The body of `countDown` typechecks because the second argument of `compareAndSet` carries annotation `χ`. The set of the witness in the constructor typechecks because newly created references are viewed as initially created with a `κ` annotation.

Our type system can validate Treiber’s algorithm for concurrent stacks, from [Goetz, Peierls, Bloch, Bowbeer, Holmes, and Lea 2005], with the following interface.

```java
class ConcurrentStack<χ> { χ Object pop();
  void push(χ Object v); }
```

When `ConcurrentStack(κ)` satisfies the usual data structure invariant, the absence of a write-read data race on the `κ` reference transmitted from the push to the pop establishes the required happens-before edge between a return of `pop` and the corresponding call of `push`. This annotated interface is validated by the following implementation. (We elide the dummy return value of `compareAndSet`.)

```java
class Node<χ> {
  final χ Object item;
  final Node<χ> next;
  Node(χ Object item, Node<χ> next) {
    this.item = item; this.next = next;
  }
}
```

```java
class ConcurrentStack<χ> {
  final Atomic<κ> top = new Atomic<κ>();
  void push(χ Object item) {
    Node<χ> oldHead;
    Node<χ> newHead;
    do { oldHead = top.get();
     newHead = new Node(item,oldHead);
    } while (!top.compareAndSet(oldHead, newHead));
  }

  χ Object pop() {
    Node<χ> oldHead;
    Node<χ> newHead;
    do { oldHead = top.get();
```
The annotation parameter of the `ConcurrentStack` (and Node) classes matches that on the items that are stored in the stack. On the other hand, the code maintains the invariant that all references to the nodes on the stack carry \( \Box \) annotation. This invariant is used to satisfy the requirement that the second arguments of the `compareToAndSet` carry \( \Box \) annotation. In order to maintain the invariant at the creation of new stack nodes in the push method, the `Node.next` field must be final.

4.4 Programs with locks

Our locks are based on Java locks, with some different conventions to simplify the semantics. Synchronization can only be performed using block-based synchronization on an instance of the Lock class. We take both Lock and the associated Condition class to be primitives; however, one can safely think of these as being implemented using atomic objects.

Locks can be used to validate the following implementation of the enriched interface for binary semaphores discussed above.

```java
class Semaphore {  
    final Lock l = new ReentrantLock();  
    final Condition open = new Condition(l);  
    final Condition acquire() {  
        return open.await(); } } } 
```

The required guarantee follows from typing. The witness field is guarded by lock \( l \), and our system ensures that all accesses to witness happen only when this lock has been acquired. The assignment to witness in acquire typechecks because of the Condition interface. In typechecking `countDown`, the Condition interface forces the argument of `release` to have \( \Box \) annotation.

The two cases of control flow for the witness returned by acquire are merged via the annotations on witness and temp and their typechecking.

5. Typing classes

We present the syntax and type rules for classes, using the annotations described in the previous section. In the next section, we describe the safety theorem: starting from a well-typed class table, (small step) evaluation of the bootstrap process cannot get stuck, and therefore, in particular, every end must always happen-after the creation of the corresponding object.

The statement of safety requires that we define well-typed classes, the evaluation relation, the bootstrap process and the stuck predicate. The proof of safety follows by defining well-typed processes, demonstrating that the bootstrap process is well-typed, and proving preservation and progress.

To simplify the semantics, we use a more restrictive language than that of the previous section. These changes are made only for conciseness of exposition: they do not alter the expressive power of the language or its type system. We make all local variables immutable and implement iteration via recursion. We make all sequencing explicit; for example return \texttt{this.f.g} must be written \texttt{val x = this.f; return x.g}.

The operational semantics follows our previous paper [Jagadeesan, Pitcher, and Riely 2010], which is in turn based on the JMM. In order to capture the properties of interest, we have enriched our previous semantics to include final fields, as well as primitive support for the Atomic and Condition classes. Unlike our previous work, we require block-structured use of Locks, which simplifies typing. Since local variables are immutable, we add a return value to these synchronization blocks, so that they can provide results to the thread without requiring the use of a mutable field.

5.1 Syntax

Let \( x \) and \( y \) range over variable names (including the reserved variable \texttt{this}), \( d \) over class names (including the reserved classes \texttt{Object}, \texttt{Atomic}, \texttt{Lock} and \texttt{Cond}), \( f \) and \( g \) over field names, and \( m \) over method names (including the reserved methods \texttt{start} and \texttt{run}).

Class types may be parameterized by annotations. Let \( \chi \) range over annotation variables and \( \gamma \) range over the ground annotations (\( \gamma ::= \Box \mid \Box \)). An annotation, \( a,b,c \), is either an annotation variable or a ground annotation (\( a,b,c ::= \chi \mid \gamma \)). Class types, \( D \), have the form \( d(\bar{a}) \). A type may be a class type or a base type (\( T,S ::= D \mid b \)). We routinely drop empty type brackets \( () \).

Let \( bV \) range over base values (including integers and the constants \texttt{unit}, \texttt{true} and \texttt{false}). The set of ground values, \( v, w, u \), includes \texttt{null} and the base values. (For running processes, the set of ground values will also include object names.) Open values include ground values and variables (\( V,W,U ::= v \mid x \)). Let \( op \) range over base value operators (such as \texttt{=}, \texttt{+, &}). We assume that \texttt{null} has meaning at all base types (zero for integers, false for booleans).

The statement language is as follows:

```plaintext
M,N ::= return V; \quad \text{(Return statement)}
| val x = new D(V); M \quad \text{(Creation statement)}
| val x = W.m(V); M \quad \text{(Method statement)}
| val x = op(V); M \quad \text{(Operator statement)}
| val x = (D)V; M \quad \text{(Cast statement)}
| end V; M \quad \text{(End statement)}
| val x = V.f; M \quad \text{(Field read statement)}
| V.f = W; M \quad \text{(Field write statement)}
| if (V) (M) else (N) \quad \text{(Conditional statement)}
| val x = sync V (N) M \quad \text{(Synchronization statement)}
```

As discussed above, synchronization statements allow a return value. All the other constructors are standard.

We often drop the keyword \texttt{val}, use complex expressions and infix notation for operators and drop instances of \texttt{“return unit”: \texttt{;}}. Thus, \( \texttt{“y = a+b+c;”} \) is sugar for \texttt{“val x = +(a,b); val y = +(x, c); return unit;”}, where \( x \) is fresh. We write \texttt{“val x = \ldots; \( M \)” as \ldots; \( \text{if} x \) does not occur free in \( M \). We write \texttt{“\uparrow V for “return V;” and \texttt{“\uparrow(V,W) for “val x = new Pair(V,W); \( \text{return x};”}} \), where \( x \) is fresh. We write \texttt{“\text{if } V \{ \text{val x = \ldots; } \( M \)” for \texttt{“if } V \{ \text{val x = \ldots; } \( M \)” else \{N\} \text{“ if } x \) does not occur free in \( M \); this notation extends to field write statements, conditional statements and sequences of statements in the obvious way.

The variable \( x \) is bound with scope \( M \) in all statements of the form \texttt{“val x = \ldots; \( M \)”}. For any syntax category, let \( fn \) return the set of free variables and let \( \text{fn} \) return the set of free names.

Class definitions have the following form:

```plaintext
D ::= class d(\chi) \{ F, M \} \quad \text{(Class declaration)}
\mathcal{A} ::= aT m(bthis; aT \{\{M\} \text{ Method declaration)}
\mathcal{F} ::= T f; \quad \text{final aD f; \text{ Field declaration)}
```
There are field declaration forms for normal, final, and lock-protected fields. Annotated fields must have object types. Annotated base types are useless since end requires an object type. Lock-protected fields have read and write guards, which are boolean combinations of field paths. When typing classes, all field paths will have the form this.f1,f2,...,fn. We elide read or write guards when uninteresting and write guarded by Φ as shorthand for rguard Φ wguard Ψ. We elide the □ annotation in variable and field declarations, and this in method declarations.

The annotation variables x are bound in the declaration class d(x)(F, M), with scope F and M. We identify syntax up to renaming and write F for the capture avoiding substitution of x for x in F, and likewise for M.

Suppose d is declared as class d(x)(F, M). Define the partial function fsyf so that fsyf(d(i,f)) = final aT if F[i][f] contains the declaration “final αT f;”, and likewise for racing and lock-protected fields. Define the partial function mtype so that mtype(d(i,m)) = bS aT → aT if M[i][m] contains the declaration aT m(bthis,f);. Define finals(d) = f if f are the final fields of class D, and similarly nonfinals.

5.2 Typing
Typing environments have the following form:

E ::= E, x : aT | E, final x : F | E, lock V

We identify environments up to reordering and repetition. Supporting judgments such as well-formed environment (E ⊢ o), type (E ⊢ T), and annotation (E ⊢ a) simply account for free and bound annotation variables. These definitions are elided.

Guard satisfaction is a propositional logic over field paths, with substitution of final paths.

E ⊢ lock F
E ⊢ F
E,x : aT ⊢ E, x,aT
E ⊢ true
E ⊢ Φ & & Ψ
E ⊢ Φ & Ψ
E ⊢ Φ & Ψ
E ⊢ Φ & Ψ

Intuitively, two guards must overlap if the sets of field paths that satisfy them must always have a non-disjoint intersection. The formal definition is a safe syntactic approximation of this intuition: Let F and G range over sets of field paths. Then any guard Φ can be written in disjunctive normal as v|Fi, where each set Fi is viewed as a conjunction of field paths. We say that Φ must overlap with Ψ if Φ = v|Fi, Ψ = v|Gj, and Fi ∩ Gj ≠ ∅ for all i, j.

Class (V), method (E ⊢ M in D) and field declarations (E ⊢ F) are as follows.

vi, x ∈Fi, x ∈Fj, x ∈Mj in d(x)
violates class d(x)(F,M, F)
E, this : bS, x : aT ⊢ M : aT
E ⊢ aT M(bthis,aT x)(M) in D
E ⊢ T
E ⊢ a a ⊢ E ⊢ D
E ⊢ T f;
E ⊢ final aDf;
E ⊢ aDf rguard Φ wguard Ψ;

These are standard definitions, with annotations added. There are three forms of field declaration, and correspondingly three type rules. The read and write guards of a lock-protected field must overlap.

Figure 1. Statement typing (rules for base types elided)

The judgments for well-typed values (E ⊢ V : aT) and statements (E ⊢ M : aT) are given in Figure 1; we have elided standard rules for base values, operators casting and conditionals.

STAT-NEW creates objects with □. The final fields of the class must be initialized by the constructor; non-final fields are set to null. STAT-END requires □ objects.

VAL-SUB-ANN1 and VAL-SUB-ANN2 define subtyping on annotations, with □ as bottom and □ as top.

In STAT-LOCK, statement N is typed under the assumption that lock V has been acquired, allowing guard satisfaction in GUARDED-READ and GUARDED-WRITE.

For final fields, the annotation on V must agree with that of the field to impart any benefit to the reader; otherwise, the reader sees □. The rules for racing fields impose no annotation restrictions on the writer, nor do they impart any annotation benefit to the reader. The rules for accessing a lock-protected field Vf require that the field and value agree in their annotation; the annotation on V is irrelevant.

6. Soundness
To validate the soundness of our system, we use a small-step reduction relation. In this section, we explain the operational semantics, define soundness, and discuss highlights of the soundness proof.

There is insufficient space to include all definitions with sufficient motivation, therefore we elide the following (a) synchronization statements (the interesting issues are already revealed by atomics); (b) discussion of auxiliary definitions used in the operational
semantics; (c) most of the additional typing rules used to prove subject reduction; and (d) all proofs. These details are included in the full version of the paper.

### 6.1 Operational semantics

Let \( p, q, s, t \) and \( \ell \) range over object names; by convention, we use name metavariables \( s, t \) for thread objects and \( f \) for atomic objects. Process syntax is as follows, where name \( p \) is bound with scope \( A \) in the process \( (vp)A \).

\[
\begin{align*}
0, u, v &::= \text{null} \mid bv \mid p \quad \text{(Ground value)} \\
M, N &::= \cdots \mid \text{val} x = \{N\} M \quad \text{(Statement)} \\
\alpha, \beta &::= (sp) \mid (sp.f=v) \mid (st\ell;j) \quad \text{(Action)} \\
A, B &::= p : D \mid t : \text{Atomic}(v; \ell) \quad \text{(Process)} \\
& \mid \text{free } p \mid \text{runnable } p \\
& \mid \alpha A \mid s[M] \mid A B \mid (vp)A \\
T &\Rightarrow A \mid \{s?p.f=v\} \Rightarrow B
\end{align*}
\]

The syntax of ground values and statements are extended to include the necessary runtime constructs. A partially executed method is written “\( \text{val} x = \{N\} M \)”, where \( N \) is the (residual of the) method body, and \( M \) is the (residual of the) calling context; that is, \( M \) is the statement that executes after the method returns.

An action indicates that the creation of an object, a write, or a synchronization has occurred. As a process executes, actions accumulate in the context of a thread. The first two process forms provide denotations for object names. The denotation of an atomic method, and composition (skim this material, coming back to it as necessary).

\[
\begin{align*}
6.1 \text{ Operational semantics} \\
\text{Let } &\bar{\sigma} = \text{act}(C, s). \text{ Let } k \text{ be the index of } s[-1] \text{ in } \bar{\sigma}. \\
&\text{• Define } C, s \text{ justifies end } p \text{ if there exists } i \text{ and } t \text{ such that } \langle \ell t p \rangle = \sigma_i \text{ and } i <_{hb} k \\
&\text{• Define } C, s \text{ justifies read } p.f=v \text{ if all of the following hold for some } i, \\
&\quad \sigma_i = \langle \ell t p.f=v \rangle \text{ for some } t \text{ (possibly equal to } s \text{), or } \sigma_i = \langle \ell t' p.f=v \rangle \text{ for some } t' \neq s \text{; and} \\
&\quad \text{for every } j \text{ such that } \sigma_j \text{ is a write action and } i <_{hb} j <_{hb} k, \text{ we have that } \text{loc}(\sigma_j) \neq \text{loc}(\sigma_i). \\
&\text{• Define } C, s \text{ justifies speculation } p.f=v \text{ if all of the following hold for some } i, \\
&\quad \sigma_i = \langle \ell t p.f=v \rangle; \\
&\quad \text{for every } j \text{ such that } \sigma_j \text{ is a write action and } i <_{hb} j <_{hb} k, \text{ we have that } \text{loc}(\sigma_j) \neq \text{loc}(\sigma_i); \text{ and} \\
&\quad \text{for every } j \text{ such that } \sigma_j \text{ is a release action and } i <_{hb} j <_{hb} k, \text{ we have that } \text{thr}(\sigma_j) \neq \text{thr}(\sigma_i). \quad \text{(continuation)}
\end{align*}
\]

Define the structural order \( A \geq B \) to be the least precongruence on processes that satisfies the axioms in Equation 6.1 (where \( \geq \) abbreviates the two axioms \( A \geq B \) and \( B \geq A \)). Define \( \equiv \) to be the kernel of \( \geq \). As usual, the structural order allows processes to move around in the syntax. It also allows actions to move outwards from a process and to commute in certain cases. The commuting rules follow the normal intuitions [Lea 2008].

The structural order \( \geq \) is defined using an order (\( \triangleright \)) on single-threaded action sequences; \( \sigma \triangleright \tau \) is defined only if \( \text{thr}(\bar{\sigma}) \cup \text{thr}(\bar{\tau}) = 1 \). Formally, \( \triangleright \) is defined to be the least precongruence on single-threaded action sequences that satisfies the following.

\[
\begin{align*}
\text{(A-NONLOCK) If } \sigma \text{ is a write or speculation, } \tau \text{ is write or speculation and } \text{loc}(\sigma) = \text{loc}(\tau) \text{ implies } \text{val}(\sigma) = \text{val}(\tau) \text{ then } \sigma \triangleright \tau. \\
\text{(A-BEGIN) If } \alpha \text{ is a begin then } \langle \ell s.p.f=v \rangle \triangleright \alpha \langle \ell s.p.f=v \rangle. \\
\text{(A-ACQUIRE) If } \alpha \text{ is write and } \beta \text{ is an acquire then } \alpha \beta \triangleright \beta \alpha. \\
\text{(A-RELEASE) If } \alpha \text{ is write and } \beta \text{ is a release then } \beta \alpha \triangleright \alpha \beta. \\
\text{(A-ABSORPTION1) If } \alpha \text{ is a write then } \alpha \beta \triangleright \alpha \alpha. \\
\text{(A-ABSORPTION2) If } \alpha, \beta \text{ and } \alpha \beta \text{ are writes to the same location then } \beta \alpha \triangleright \alpha \beta. \\
\text{(A-ABSORPTION3) If } \alpha, \beta \text{ and } \beta' \text{ are writes to the same location then } \beta' \alpha \triangleright \beta \alpha. \\
\end{align*}
\]

Define the reduction relation \( A \rightarrow B \) to be the least relation satisfying the rules and axioms in Equations 6.1. Define \( \rightarrow \) to be the reflexive and transitive closure of \( \rightarrow \). In the figure, \( \mathcal{B} \mathcal{C} \) is the set of base values. We rely on many other standard auxiliary definitions.

---

1 In classes and methods, the parameters are bound. In statements, bound variables are preceded by \( \text{val} \). In processes, bound names are preceded by \( v \). There are no binders in contexts. We identify syntax up to renaming of bound variables and names and write \( M[[/s]] \) for the capture avoiding substitution of \( v \) for \( x \) in \( M \). We assume similar notation for substitution of names for names and for substitution over other syntax categories.

Suppose \( d \) is declared as class \( d(\bar{C}) \in \mathcal{C} \). Define the partial function \( \text{type} \) so that \( \text{type}(d(\bar{C})) = \tau \) if \( d(\bar{C}) \in \mathcal{C} \) contains the declaration \( \tau_1 \ll a T ; \tau_2 \) and \( \tau_1 \neq \tau_2 \). Define the type declaration \( \text{decl} \) so that \( \text{decl}(d(\bar{C})) = M \) if \( d(\bar{C}) \in \mathcal{C} \) contains the declaration \( a T M \).
Many of the reduction rules are standard. We very briefly discuss some of the exceptions.

Atomic objects maintain a shared global state, enforcing sequential consistency. The state includes an integer, which is odd if the last operation was a set. These objects generate synchronization actions, which define the synchronizes-with relation. Even actions may be seen as an get/accquire and odd actions a set/release. In the formal development, we take it that can returns null if the comparison fails, and the prior object otherwise.

As in Java, the reserved method \texttt{start} starts method run under the thread identity of the receiving object; this is a synchronization event enforced using a fresh "dummy" lock.

R-FIELD-WRITE describes field writes in a relaxed memory model, so the field writes become actions that float into the evaluation context. Field reads, R-FIELD-READ, may take any value that is justified by the evaluation context. R-END removes an end if the creation of the object reference is in a happens-before relationship; this is used in the following sections on typing.

Speculation can occur at any point, using R-SPECULATION-OPEN. The initial branch has guard \( T \), indicating that this branch may make no additional assumptions. The final branch has a speculative action as its guard that can be used to justify reads. The context rules permit each branch of speculation evolve independently. Results from an active speculation can only leak to the outside world via S-PREFIX-SPECULATION if all branches produce the same action. This is significant, since only actions that manage to make it outside of a speculation may be used to finalize via R-SPECULATION-CLOSE.

6.2 Defining soundness

Soundness is defined in terms of \textit{stack} threads, which are defined in terms of evaluation contexts, defined as follows.

\[
\begin{align*}
E &::= [-] \mid \text{val} \ x = \{\text{val}\} M \\
C &::= [-] \mid \text{stmt} A \mid \text{stmt} C \mid \text{val} P \mid \text{var} C \
| \text{stmt} C 
\end{align*}
\]

(Statement ctxt)

\[
\begin{align*}
T \Rightarrow C \mid \langle s \text{?} p. f = w \rangle \Rightarrow A \\
T \Rightarrow A \mid \langle s \text{?} p. f = w \rangle \Rightarrow C
\end{align*}
\]

(Process ctxt)

Roughly, a stuck thread is a non-terminated thread that cannot reduce. As discussed by Igarashi, Pierce, and Waldner [2001], Java allows runtime type exceptions which are not considered a failure.
of the static type system. Here we also have null pointer exceptions\(^2\).

**Definition 1 (Stuck thread).** Thread \(s\) is stuck in \(A\) if there exists \(\mathbb{C}[s]\mathbb{E}[M] = A\) such that none of the following hold.

1. \(s\) has terminated: \(M\) is a return statement and \(\mathbb{E} = \bot\);
2. \(s\) can reduce: \(\mathbb{C}[s]\mathbb{E}[M] \Rightarrow \mathbb{C}'[s]\mathbb{E}'[M']\) for some \(\mathbb{C}', M';\)
3. \(s\) has had a class cast exception: \(M = \text{val}x = (D)v\); \(M'\) and \(\mathbb{C}\) does not contain subterm \(v; D\); or
4. \(s\) has had a null pointer exception: \(M\) is a method, end, read, write, conditional or synchronization statement with target null.

The stuck processes of most interest come about because of the premise of the reduction rule for end statements. In order for end \(p\) to reduce, it must be the case the action \((t,p)\) happens-before, for some \(t\). This is a global property of the process, tracking data flow.

**Theorem 2 (Soundness).** Suppose the class table is well typed and method \(\text{Main} \cdot \text{main}()\) is defined. Suppose further that \(A\) is a speculation-free process that is reachable from the bootstrap process, that is

\[
(\text{vm} : \text{m} \cdot \text{Main} | \text{m} \cdot \text{n} \cdot \text{main}();) \rightarrow \cdots \rightarrow A.
\]

Then \(A\) does not contain a stuck thread.

We restrict attention to speculation-free processes, because threads within the final branch of a “bad” speculation may get stuck. Typing ensures that such speculations can never be finalized.

### 7. Typing processes

The proof of soundness follows standard lines: We provide typing rules for processes and prove subject reduction and progress. The details of the proof are too long to include here. Indeed, even the invariant maintained by running time is too long to fully describe here. In this section, we describe some of the most interesting typing rules for processes, which embody intermediate states of computation. We first discuss the rules for “standard” processes, then for speculation.

In order to type processes, we introduce *effects* and extend environments as follows.

\[
\xi, \xi' := \emptyset[p]
\]

\[
E := \cdots \ell, E, s; \text{lock} \ell | E, s; \text{returns} a D | E, \xi'| s, \xi; | E, \xi; \ell, \xi' | \text{onacq} F \xi
\]

The effect \(\emptyset[p]\) indicates that \((t, p)\) has occurred, for some \(t\). Effects are associated with threads \((s; \xi)\) and synchronization objects, or locks \((\ell; \xi)\). Naked effects \((\xi)\), used in statement typing, are associated with the “local” thread. The environment \(s; \text{lock} \ell\) indicates that \(s\) holds lock \(\ell\). The onacq environment is used to type bad speculations, as discussed in Section 7.1.

Define \(\text{locks}(s, E) = \{\text{lock} V | s; \text{lock} V \in E\}\) and \(\text{effects}(s, E) = \{\xi | s; \xi\} \in E\). Define \(E \vdash \phi\) to mean \(E, \text{effects}(s, E), \text{locks}(s, E) \vdash \phi\), and similarly for \(E \vdash V : a T\) and \(E \vdash M : a T\). Let \(Z\) range over sets of effects.

The new rules for values and the key rules for processes \((E \vdash A)\) and actions \((E \vdash \sigma \triangleright E')\) are given in Figure 3. The proof of soundness also uses *well-formedness*, which is defined in the full version of this paper. Well-formedness is trivially true of the bootstrap process and is preserved by structural order and reduction. Well-formedness eliminates nonsense processes, such as \(s[M] \triangleright s[V]\) and \((t, p, f = v) \triangleright (t, p, f = w) A\), where \(f\) is a final field.

The action sequence records the entire history of the running process. The type rules can compute the happens-before relation as it moves through the action context. The same is true for the locks held at any point. Subject reduction uses the fact that such sequences are stable up to single-threaded action reordering.

**ACT-ACQUIRE** adds a lock to the environment and copies effects from locks to threads. **ACT-RELEASE** removes a lock from the environment and copies effects from threads to locks. (We discuss onacq in in Section 7.1.)

Using **ACT-BEGIN** and **ACT-CONDITION**, the effect \(\emptyset[p]\) is available to \(A\) in \((t, p) A\). The VAL-OBJECT rules subsequently require that \(\emptyset[p]\) be present in order to type \(p\) at \(\mathbb{E}\).

**ACT-GUARDED-WRITE** directly parallels **STAT-GUARDED-WRITE**, ensuring that appropriate locks and effects are in scope. The locks are only necessary after the fields of \(p\) are initialized and thus \(t : [p]\) is in scope, for some \(t\). This rule is preserved by action reordering: Although write actions can commute with some write actions, the rules require that locks can only expand their scope.

#### 7.1 Speculation

Several of the typing rules for speculation are given in Figure 4.
The speculation \( T \Rightarrow A \langle s ? p.f=q \rangle \Rightarrow B \) allows the justifying branch \( A \) and final branch \( B \) to proceed independently, and also allows the final branch to speculate on a future write \( \langle s ? p.f=q \rangle \). The operational semantics does not constrain the speculated write, so the speculated write may not be well-typed. Consequently, a read of the speculated write by the final branch process \( B \) can cause the resulting final branch to fail to type. Moreover, the speculation can reduce to the final branch if the speculated write is justified by the speculation’s context. Thus it is natural to ask whether a process with ill-typing confined to the final branch of a speculation can reduce to a process with ill-typing outside speculation.

There are two classes of typing errors that may arise. First, shape errors may arise when a speculated write refers to an object with ill-typing confined to the final branch of a speculation can cause the resulting final branch to fail to type. Moreover, the speculation can reduce to the final branch if the speculated write is justified by the speculation’s context. Thus it is natural to ask whether a process with ill-typing confined to the final branch of a speculation can reduce to a process with ill-typing outside speculation.

Such shape errors are permitted in final branches of speculations. However, such speculations are *unfinalizable* since a well-typed justifying branch, which does not see the speculation, cannot produce the necessary write to justify the speculation. The use of the justifying branch of a speculation as a filter is critical to prevent ill-typed actions leaking from the final branch of a speculation.

Type errors may also arise in the final branch of a speculation due to a misuse of effects, but such speculations are also unfinalizable. Unlike shape errors, it is not immediately evident that an effect error has occurred. Our approach with effects, therefore, is to saturate the final branch with sufficient effects to ensure that no type errors arise, while ensuring that all of the imparted effects can be justified when the speculation is finalized.

A read of the speculated write may convey information about happens-before relationships via \( \square \)-annotated types that is not justified by the process context. For example, a speculated write \( \langle s ? p.f=q \rangle \) may be read in the final branch of a speculation in the absence of a happens-before relationship with \( [q] \) (thus preventing a write action \( \langle s ? p.f=q \rangle \) being typeable at the point of the speculation).

The type system accommodates potential memory-effect errors by associating *specified effects* \( s \) with speculated writes \( \langle s ? p.f=q \rangle \), whenever the corresponding write \( \langle s ? p.f=q \rangle \) would convey a memory effect. This allows the final branches of speculations to be typed even if reduction reads speculated writes that do not yet convey happens-before relationships. Speculated effects can be discharged when their associated speculations are finalized because the justifying writes for the speculation appear outside the speculation and thus justify the speculated memory effects.

Typing for speculation on a final field does not require locks or latent effects. Instead, as with reads of final fields, it is significant whether a thread has the effect \( [p] \) given a speculated write \( \langle s ? p.f=q \rangle \). \( \text{ACT-\square-FINAL-SPECULATION} \) generates an environment extension with an effect \( [q] \) for all threads that prove \( [p] \). If the speculation on a final field’s write is finalized, the use of \( \text{ACT-\square-FINAL-SPECULATION} \) can be discharged immediately in the resulting final branch.

It is worth emphasizing that the additional effects provided by speculative fields are only necessary for speculations that cannot be finalized. Our semantics has been designed to allow as many speculations as possible.

Alternatively, one could attempt to disallow "useless" speculations, which either (1) cannot be finalized or (2) add nothing to the execution. For example, one can show that speculation with a value of the wrong shape is useless, for the first reason: speculating on a lock-protected field is useless, for the second reason. It seems difficult, however to prohibit useless speculation on final fields, which communicate effects. In addition, our approach can be adapted to richer effect languages.

### 7.2 Typing examples

The environments used to type processes have some shared information, common to all threads, such as the type of an object reference \( p \). In addition, environments record local information including effects held on a per thread basis. Consequently, each thread may have a different view of the effects associated with an object.

For example, consider the following process (we leave names free, and omit uninteresting local variable bindings and continuation terms). Let \( p:C5 \).

```java
class C5 {final Lock l; Object f guard this.l;}

s[val x=ev Object(); sync p.l(p.f=x;)] |
t[sync p.l(val y=p.f; end y;)]
```

After the constructor, thread \( s \) will be typed in an environment including \( t: \square \text{Object} \). In the presence of suitable auxiliary processes such as "free \( q \)" the above process may reduce to the following, where \( p:C5, q:\text{Object} \).

```java

\( \langle s?q \rangle (s[sync p.l(val f=q;)] | t[sync p.l(val y=p.f; end y;)])
```

The environment for typing the threads includes \( p:C5, q:\text{Object} \).

\( s: \[ q \] \). The localized effect \( s: [q] \) represents the fact that code running at thread \( s \) has a happens-before relationship with \( q \)’s constructor. This effect is introduced into the environment by \( \text{PROC-\square-BEGIN} \)’s use of \( \text{ACT-\square-BEGIN} \) when typing the action \( \langle s?p.f=q \rangle \). The same effect is required to justify the typing of \( q \) as \( \square \text{Object} \) by \( \text{VAL-\square-\text{OBJECT}} \). Without the effect \( s: [q] \), uses of \( q \) could only be typed as \( \square \text{Object} \) via \( \text{VAL-\square-\text{OBJECT}} \).

The effect \( s: [q] \) cannot be used directly within thread \( t \), but is communicated in the statics (as in the dynamics) via memory synchronization actions such as lock acquisition and release. To see this, consider a further three reductions by thread \( s \) above. We assume that a write of \( k \) to \( p \).1 by some thread is visible. (We apply structural reordering for readability and again omit auxiliary processes. We assume that lock \( k \) has not yet been used before). Let \( p:C5, q:\text{Object}, k:\text{Lock} \).

```java

\( \langle s?q \rangle (s[t:k:0] (s[t:f=q] (s[t:k:42]) \langle s[l] \rangle t[sync p.l(val y=p.f; end y;)])
```

The environment records that lock \( k \) is held by thread \( s \) between the actions \( \langle s[t:k:0] \rangle \) and \( \langle s[t:k:42] \rangle \). In particular, justifies the write \( \langle s[t:f=q] \rangle \) to a \( \square \) field guarded by \( k \). Moreover, the assignment of \( q \) to \( a \) \( \square \) field \( \langle s[t:f=q] \rangle \) requires the \( s: [q] \) effect in the environment at the point of the write, mirroring the assignment typing in terms described above.

Further reductions lead to the following process where the lock \( k \) has been acquired by \( t \), and \( q \) has been read from \( p.f \) then substituted for \( z \). Let \( p:C5, q:\text{Object}, k:\text{Lock} \).

```java

\( \langle s?q \rangle (s[t:k:0] (s[t:f=q] (s[t:k:42]) (t[tk:27]) (s[l] t[\langle \text{end q;} k:release(); \rangle])
```

Here the typing of \( q \) via \( \text{STAT-END} \) requires that \( q \) be typed as \( \square \text{Object} \) within thread \( t \), and, as before, this requires that the environment provides \( [q] \) for thread \( t \). The lock release \( \langle s[t:k:42] \rangle \) and lock acquire \( \langle t[tk:27] \rangle \) achieve this by transferring the effects known to thread \( s \) \( (s: [q]) \) to lock \( k \) \( (k: [q]) \) at the release of \( k \) by \( s \), and transferring the effects for lock \( k \) to thread \( t \) \( (t: [q]) \) at the acquisition of \( k \) by thread \( t \). Thus we type the threads in an environment including \( t: [q] \). In this way, the type system calculates communication of effects between threads following the memory synchronization actions.
The typing of final fields necessarily differs from the typing of fields guarded by locks because there are no memory synchronization actions used when reading a final field. The key observation in reading a final field is that the object reference through which the final field is read must be $\exists$ annotated. For example, in typing the following process, $\langle \text{end } x \rangle$ is only permitted if $x : \exists \text{ Object}$, and this in turns requires $\exists : [p]$ in the typing environment for the process. Let $p : C6$.

```
class C6 {final @Object f;}
  s[val x=p.f; end x;]
```

This is expressed in the rule $\text{STAT-FINAL-READ}$.

Although writes to final fields may migrate via the structural order, this invariant is sufficient because writes to final fields always happen-before the begin representing construction of the object. The $\exists$ nature of a value $q$ read from a final field $p.f$ then follows from transitivity of the happens-before relation between: (1) the construction of $q$ and the write to $p.f$; (2) the write to $p.f$ and the begin for $p$; (3) the begin for $p$ and its receipt at $\exists$ type; and (4) receipt of $p$ and the read $p.f$. This property demonstrates how happens-before relationships can be established for lockless traversal of immutable portions of data structures whenever a happens-before relationship is established with the construction of the initial object reference.

### 8. Related work and conclusion

We have already cited several related pieces of work in context in the paper. We refer the reader to [Steinke and Nutt 2004] and [Adv and Ghar 1996] for surveys of memory models for hardware architectures. Saraswat [2004] provides a framework for operational semantics with relaxed memory models for typed languages. Our prior work [Jagadeesan, Pitcher, and Riely 2010] uses the techniques of Cenciarelli, Knapp, and Sibilio [2007] and Boudol and Petri [2009] to accommodate the full expressiveness of the JMM.

The $\exists$ annotation uses objects to witness side-effects [Terauchi and Aiken 2008]. However, in contrast to the large research on data races, we only enforce the absence of read-write data races on our witness objects and allow write-write data races. Happens-before plays a key role in the static analysis of programs, e.g., vector-clock algorithms for race detection [Flanagan and Freund 2009] or the characterization of execution paths [Burckhardt, Kothari, Musuvathi, and Nagarackatte 2010].

There has also been recent work on static analysis for weak memory models [Ferrara 2008; Miné 2011; Alglave, Kroening, Lugton, Nimal, and Tautschnig 2011] and provably correct compilers targeting weak memory Sevcik, Vafeiadis, Nardelli, Jagannathan, and Sewell [2011].

The study of relaxed memory models has hitherto focused on the dynamic semantics of programs. This paper has focused on static semantics. We have argued that correspondence assertions are the correct formalism for reasoning about happens-before relations, and that these correspondences can be specified and validated by enriching interfaces with notation for memory effects. Our techniques apply to common data structures from java.util.concurrent, including both lock-based and non-blocking implementations.

Our analysis intentionally includes only a small fragment of the type system of Abadi, Flanagan, and Freund [2006] for data race detection. Fully integrating other techniques used there (such as existential types in the form of ghost variables) and other existing techniques, including analyses based on locality, confinement, linearity, monotonicity, and effects is a topic of future research.

### References


A. Igarashi, B. C. Pierce, and P. Wadler. Featherweight Java: a minimal core calculus for Java and GJ. ACM TOPLAS, 23(3):396–450, 2001.


A. Syntax
Figure 5 collects the syntactic definitions from the text. The syntax differs from that of the body text only in that we have introduced speculative actions and used these to define speculative processes. We have also introduced a category for actions. For bookkeeping purposes, we also annotate partially executed method bodies and synchronization blocks with the identity of the thread in which they occur.

Figure 6 includes the additional reduction rules for locks and conditions.

B. Well-formed processes
Define \( \text{procs} \), mapping contexts to processes, as follows.
\[
\begin{align*}
\text{procs}(\_\_\_) & = \emptyset \\
\text{procs}(\alpha \ C) & = \text{procs}(C) \\
\text{procs}(\nu q \ C) & = \text{procs}(C) \\
\text{procs}(\xi A) & = A \cup \text{procs}(C) \\
\text{procs}(A \phi \Rightarrow B) & = \text{procs}(C) \cup \text{procs}(C) \cup \text{procs}(A) \cup \text{procs}(\alpha \ C) \cup \text{procs}(A) \cup \text{procs}(B)
\end{align*}
\]
A process \( A \) is well-formed if it satisfies the following.
1. If \( A \in C[\text{free}], C[\text{runnable}] \) or \( C[s[M]] \) then the following hold.
   a. \( C \) is contiguous.
   b. There is no \( \text{free} \), \( \text{runnable} \) or \( s[N] \) in \( \text{procs}(C) \).
   c. At most one of \( \text{free} \), \( \text{runnable} \) or \( t[N] \) occurs in \( \text{procs}(C) \), for any \( t \).
   d. At most one \( p:D \) occurs in \( \text{procs}(C) \), for any \( p \).
   e. If there is an action in \( \text{act}(C,s) \) with source \( t \neq s \), then \( t[N] \) occurs in \( \text{procs}(C) \).
   f. If \( p \) occurs in \( C \) or \( M \), then \( p:D \) occurs in \( \text{procs}(C) \).
   g. If \( p:D \) occurs in \( \text{procs}(C) \) and \( f \in \text{finals}(D) \), then there exists exactly one write action to \( p.f \) in \( \text{act}(C,s) \).
2. Let \( \text{thrds} \) and \( \text{objs} \) return the names of the potential threads and the objects of a process, respectively. For example \( \text{thrds}(\text{free}) = \text{thrds}(\text{runnable}) = \text{thrds}(s[M]) = \{s\} \) and \( \text{objs}(p:D) = \text{objs}(p:\text{atomic}(v)) = \text{objs}(p:\text{Lock}(G,s)) = \{p\} \).
   a. In any subprocess \( B[B'] \), we have that \( \text{obj}(B) \cap \text{obj}(B') = \emptyset \), \( \text{thrds}(B) \cap \text{thrds}(B') = \emptyset, \text{thrds}(\alpha) \cap \text{thrds}(B) = \emptyset \).
   b. In any subprocess \( aB \), we have that \( \text{thrds}(\alpha) \cap \text{thrds}(B) = \emptyset \).
   c. In any subprocess \( T \Rightarrow B[B' \phi \Rightarrow B'] \), we have that \( \text{thrds}(\phi) \cap \text{thrds}(B) \cap \text{thrds}(B') \).

Well-formedness is trivially true of the bootstrap process and is preserved by structural order and reduction.

**Lemma 3. The bootstrap process is well-formed.**
**Proof.** Trivial. \(\square\)

**Lemma 4.** (a) Suppose \( A \) is well-formed and \( A \geq A' \) then \( A' \) is well-formed.
(b) Suppose \( A \) is well-formed and \( A \rightarrow A' \) then \( A' \) is well-formed.
**Proof.** (a) follows by induction on the proof that \( A \geq A' \). (b) follows by induction on the proof that \( A \rightarrow A' \). \(\square\)

C. Typing details
The reserved methods \texttt{start} and \texttt{run} have type “\( \exists \theta \rightarrow \square \text{Unit} \).”

The reserved classes have the following signatures. Recall that we elide “\( \square \text{thin} \)” from method interfaces. In the formal development, we take it that \( \text{cas} \) returns null if the comparison fails, and the prior object otherwise.

```plaintext
class Object { }
class Atomic(x) {
  x Object avai();
  void set(x Object v);
  x Object cas(Object oldv, x Object newv);
}
class Lock { }
class Cond(x) {
  x Object get();
  void signal(x Object);
}
```

\(\chi\) Object avai();
void signal(\(\chi\) Object);
}

We now present the judgments, divided into five groups:

---

**Judgments for satisfaction.**
\[
\begin{align*}
E \models \Phi & \quad \text{(Guard satisfaction)} \\
E \models \zeta & \quad \text{(Effect satisfaction)}
\end{align*}
\]

**Judgments for statics.**
\[
\begin{align*}
\models \phi & \quad \text{(Ok class declaration)} \\
E \vdash \Phi in D & \quad \text{(Ok method declaration)} \\
E \vdash \Phi & \quad \text{(Ok field declaration)} \\
E \vdash V : aT & \quad \text{(Ok value)} \\
E \vdash M : aT \Rightarrow Z & \quad \text{(Ok statement)}
\end{align*}
\]

**Judgments for dynamics.**
\[
\begin{align*}
E \models A & \quad \text{(Ok process)} \\
E \sigma \models E' & \quad \text{(Ok action)}
\end{align*}
\]

**Supporting judgements.**
\[
\begin{align*}
E \models \phi & \quad \text{(Ok environment)} \\
E \models T & \quad \text{(Ok type)} \\
E \vdash \sigma & \quad \text{(Ok annotation)} \\
E \vdash F : T & \quad \text{(Ok path)} \\
E \vdash \Phi & \quad \text{(Ok guard)} \\
E \vdash \zeta & \quad \text{(Ok effect)}
\end{align*}
\]

**Judgments for contexts.**
\[
E \models \zeta & \quad \text{(Ok context)}
\]

We identify environments up to reordering.

---

**C.1 Judgements for satisfaction**

**Judgment** \( E \models \Phi \).
\[
\begin{align*}
E \models \text{lock} F & \quad E \models \Phi & \quad E \models \Psi & \quad E \models \Phi \cap \Psi & \quad E \models \Phi \cap \Psi
\end{align*}
\]

**E \models \text{false} \Rightarrow \Phi, E \models \text{true} \Rightarrow \Phi, E \models \Phi \cup \Psi \Rightarrow \Phi, E \models \Phi \cap \Psi, E \models \Phi \cap \Psi
\]

**E \models \text{false} \Rightarrow \Phi, E \models \text{true} \Rightarrow \Phi, E \models \Phi \cup \Psi \Rightarrow \Phi, E \models \Phi \cap \Psi
\]

**E \models \text{false} \Rightarrow \Phi, E \models \text{true} \Rightarrow \Phi, E \models \Phi \cup \Psi \Rightarrow \Phi, E \models \Phi \cap \Psi
\]

---

**C.2 Judgements for statics**

Let \( F \) and \( G \) range over sets of field paths. Then any guard \( \Phi \) can be written in disjunctive normal as \( \bigvee_i F_i \), where each set \( F_i \) is viewed as a disjunction of field paths. We say that \( \Phi \) must overlap with \( \Psi \) if \( \Phi = \bigvee_i F_i, \Psi = \bigvee_j G_j \), and \( F_i \cap G_j \neq \emptyset \) for all \( i, j \).

**Judgments** \( \models \phi \) and \( E \vdash \Phi \in D \).
\[
\begin{align*}
\forall i. \forall j. \forall x. E \models \text{false} \Rightarrow d(F_i(x)) & \quad E \models \text{true} \Rightarrow d(F_i(x)) & \quad E \models \text{false} \Rightarrow d(F_i(x)) & \quad E \models \text{true} \Rightarrow d(F_i(x)) & \quad E \models \text{false} \Rightarrow d(F_i(x)) & \quad E \models \text{true} \Rightarrow d(F_i(x))
\end{align*}
\]

**Judgment** \( E \vdash \Phi.\)
\[
\begin{align*}
E \vdash T & \quad E \vdash a & \quad E \vdash D & \quad E \vdash \Phi.\Psi & \quad E \vdash D : \text{false} \Rightarrow \Phi.\Psi & \quad E \vdash D : \text{false} \Rightarrow \Phi.\Psi & \quad E \vdash D : \text{false} \Rightarrow \Phi.\Psi & \quad E \vdash D : \text{false} \Rightarrow \Phi.\Psi
\end{align*}
\]

**Judgment** \( E \vdash V : aT.\)
\[
\begin{align*}
E \vdash a & \quad E \vdash V : aT & \quad E \vdash V : aT & \quad E \vdash V : aT & \quad E \vdash V : aT & \quad E \vdash V : aT
\end{align*}
\]

---

**Val-base**
\[
\begin{align*}
E \vdash \text{true} & \quad \text{bt} \models \text{true} & \quad E \vdash \text{false} & \quad \text{bt} \models \text{false} & \quad E \vdash \text{false} & \quad \text{bt} \models \text{false} & \quad E \vdash \text{false} & \quad \text{bt} \models \text{false} & \quad E \vdash \text{false} & \quad \text{bt} \models \text{false} & \quad E \vdash \text{false} & \quad \text{bt} \models \text{false}
\end{align*}
\]
\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Syntactic categories}
\end{figure}
Figure 6. Reduction rules for locks and conditions (A → B)

C.3 Judgements for dynamics

Define \( \text{locks}(s, E) = \{ \text{lock } V \mid s \text{ lock } V \in E \} \) and \( \text{effects}(s, E) = \{ \zeta \mid s, \zeta \in E \} \). Define \( E \models_\Phi \Phi \) to mean \( E, \text{effects}(s, E), \text{locks}(s, E) \models_\Phi \Phi \), and similarly for \( E \models_\Gamma V : aT \) and \( E \models_\Gamma M : aT \).

Judgment \( E \vdash A \).

```

```

```
C.5 Judgments for contexts

The following judgment is used in proofs to connect contexts and environments. See Appendix D.3.

\[ E \vdash C \triangleright E' \]

\[
\begin{align*}
\text{(context-par-1)} & \quad E \vdash A \quad E \vdash C \triangleright E' \\
\text{(context-par-2)} & \quad E \vdash A \quad E \vdash C \triangleright E' \\
\text{(context-nv)} & \quad E \vdash p: D \quad p \text{ returns } \Box \text{ Unit } \vdash C \triangleright E' \\
\text{(context-action)} & \quad E \vdash \alpha \triangleright E' \quad E' \vdash E'' \\
\text{(context-speculation-initial)} & \quad E \vdash C \triangleright E' \\
\text{(context-speculation-final-shape-error)} & \quad E \vdash A \quad E \vdash \phi \triangleright E' \quad E' \vdash B \\
\text{(context-speculation-final-shape-error)} & \quad E \vdash \alpha C \triangleright E' \\
\text{(context-initial-shape-error)} & \quad E \vdash p: D \quad E \vdash C \triangleright E' \\
\text{(context-speculation-final-shape-error)} & \quad E \vdash A \quad E \vdash \phi \triangleright E' \quad E' \vdash C \triangleright E'' \\
\end{align*}
\]

D. Proofs

Soundness follows from progress and preservation. Recall the soundness theorem.

**Definition (1) (Stuck thread).** Thread s is stuck in A if there exists C \[ s(E[M]) \] in A such that none of the following hold:

1. s has terminated: M is a return statement and E = \[ \llbracket \cdot \rrbracket \];
2. s can reduce: C \[ s(E[M]) \] \rightarrow C \[ s(M') \] for some C', M';
3. s has had a class cast exception: M = val x = (D)v; M' and C does not contain subterm v; or
4. s has had a null pointer exception: M is a method, end, read, write, conditional or synchronization statement with target null.
5. s is waiting for a lock, i.e., M is a synchronization statement or call to Cond.await;
Theorem (2) (Soundness). Suppose the state class table is well typed and method \texttt{Main.main()} is defined. Suppose further that \( A \) is a speculation-free process that is reachable from the bootstrap process, that is

\[
\forall m : \text{Main} | m.\text{main}(); \rightarrow \cdots \rightarrow A.
\]

Then \( A \) does not contain a stuck thread.

**Proof.** By inspection, the bootstrap process is well-formed and well-typed under the empty environment.

As shown in Appendix B, well-formedness is preserved by reduction.

By Proposition 30, proved in Appendix D.7 below, well-typedness is preserved by reduction.

By Proposition 21, proved in Appendix D.5 below, no well-formed and well-typed processes may contain a stuck thread.

To prove progress and preservation, we first introduce the notion of a trap, i.e., a context that contains a bad speculation. In Appendix D.3, we provide composition/decomposition results enabling us to go back and forth between contexts and processes. We relate the dynamic and static definitions of happens-in before Appendix D.4. We state progress in Appendix D.5 and give a brief outline of the proof. We prove preservation in two parts, treating structural order in Appendix D.6 and reduction in Appendix D.7.

### D.1 Traps

Define \( \text{lockenv}(E, C) \) as follows.

\[
\text{lockenv}(E, [\ell]) = E
\]

\[
\text{lockenv}(E, C(A)) = \text{lockenv}(E, C)
\]

\[
\text{lockenv}(E, \ell(A)) = \text{lockenv}(E, C)
\]

\[
\text{lockenv}(E, \alpha) = \text{lockenv}(E, C)
\]

if \( \alpha \) nonlock

\[
\text{lockenv}(E, \ell(t); j) = \text{lockenv}(E, \ell(x; \text{lock } t), C)
\]

if \( j \) even

\[
\text{lockenv}(E, \ell(t); j) = \text{lockenv}(E, \ell(x; \text{lock } t), C)
\]

if \( j \) odd

Define \( \text{lockenv}(C) = \text{lockenv}(\emptyset, C) \). Define \( C \vdash \tau \Phi \) as shorthand for \( \text{lockenv}(C) \vdash \tau \Phi \).

**Definition 5 (Trap).** A context \( C \) is a shape trap if \( C \supseteq p:D \),

\[
C = C' \vdash [\top \Rightarrow A \vdash \langle \pi t p, f = w \rangle \Rightarrow C'']
\]

and either

* \( \{\text{fype}(\ell(D,f)) = \ldots bt \ldots \text{ and } \neg (v \text{ inhabits } bt) \} \), or
* \( \{\text{fype}(\ell(D,f)) = \ldots D' \ldots \text{ and } \neg (E \supseteq v:D') \} \).

A context \( C = C' \vdash [\top \Rightarrow A \vdash \langle \pi t p, f = w \rangle \Rightarrow C''] \) is a trap if \( C \supseteq p:D \) and either it is a shape trap or

* \( \{\text{fype}(\ell(D,f)) = \cdots T \text{ wr-guard } \Psi \text{ and there exists } D'[\Psi''] = C'' \text{ such that } (a) \neg (C'[\Psi']) \vdash \Psi \text{, and } (b) \text{ for every } D'[\Psi'''] \supseteq D, \neg (C''[\Psi''']) \text{ justifies speculation } p,f = w \} \).

The last possibility indicates that \( s \) has given up the locks required to justify the speculation and has not done the necessary write. A shape trap can always be typed using \( \text{PROC-SPECULATION-SHAPE-ERROR} \).

**Lemma 6 (Trap).** Suppose \( C \) is a trap, \( \emptyset \vdash C\langle v[M] \rangle \vdash C'[v[M']] \). Then, \( C'' \vdash C \) is a trap.

**Proof.** A trap can only be removed by \( \text{SPECLCATION-CLOSE} \). Assume for the sake of contradiction that there exists a trap that can be removed by \( \text{SPECLCATION-CLOSE} \). Thus, we have a speculation of the form

\[
\top \Rightarrow A \vdash \langle \pi t p, f = w \rangle \Rightarrow C''
\]

and that (1) \( C'\vdash s \text{ justifies speculation } p,f = w \). As (1) holds, we know that there exists some \( i \) such that \( \sigma_i = (\pi t p, f = w) \). However, this would require (2) \( \{\text{fype}(\ell(D,f)) = \ldots bt \ldots \text{ and } v \text{ inhabits } bt \} \) or \( \{\text{fype}(\ell(D,f)) = \ldots D' \ldots \text{ and } E \supseteq v:D' \} \), which contradicts the first two premises of the definition of a trap.

For the third possibility, note that by typing rules for actions, \( s \) must hold the necessary locks to produce a justifying write action, \( s \) cannot acquire the locks to do so after the speculation, since in this case the write cannot leave the speculation (synchronization actions do not commute with speculation).

So \( s \) must have acquired the necessary locks before the speculation, and must perform the write before releasing the locks, as \( s \) cannot acquire them again and subsequently justify the speculation. But this case is exactly prohibited by the third clause of the definition.

Therefore, the original assumption is false, and the lemma holds.

**D.2 Weakening/Strengthening**

The proofs of the following lemmas follow standard lines.

**Lemma 7 (Weakening).** If \( E \vdash A, E', \sigma \vdash E' \), and \( E' \vdash \sigma \vdash E' \), then \( E, E' \vdash \sigma \vdash E' \). Similarly for other judgments.

**Corollary 8.** (a) Assume \( p \notin \text{fn}(\sigma) \) and \( E \vdash \sigma \vdash E' \). Then, \( E, p:D \vdash \sigma \vdash E', p:D \).

(b) Assume \( E \vdash A \) and \( E \vdash \sigma \vdash E' \) such that \( A \) doesn’t have any synchronization actions that overlap with \( \alpha \). Then, \( E' \vdash \alpha \).

(c) Assume \( E \vdash A \) and \( E \vdash \sigma \vdash E' \). Then, \( E' \vdash \sigma \).

**Lemma 9 (Strengthening).** If \( E, x:bS \vdash M : aT \), and \( x \notin M \), then \( E \vdash M : aT \).

If \( E, p:D \vdash \sigma \vdash E', p:D, E'' \). Then, \( E \vdash \sigma \vdash E', E'' \).

**Corollary 10.** Assume \( p \notin \text{fn}(\sigma) \) and \( E, p:D \vdash \sigma \vdash E', p:D, E'' \). Then, \( E \vdash \sigma \vdash E', E'' \).

**Definition 11 (Sensible extension).** Suppose environment \( E \) is well formed. We define \( E' \) to be a sensible extension of \( E \) if \( E' \) is well formed and there exists \( s, t, f, E' \) such that:

\[
E = E; A; \text{lock } t_1, \ldots, t_n; \text{lock } e_1 \quad \text{and} \quad E' = E'.
\]

**Lemma 12 (Actions preserve wellformed environments).** Suppose \( E \vdash \sigma \vdash E' \) and \( E \vdash \sigma \vdash E' \). Then of \( E \vdash \sigma \).

**Proof.** By induction on \( E \vdash \sigma \).

**Lemma 13 (Actions sensibly extend environments).** Suppose \( E \vdash \sigma \vdash E' \). \( E' \) is a sensible extension of \( E \).

**Proof.** We prove this by analysis of the \( E \vdash \sigma \vdash E' \) judgments. We first note that for \( \text{ACT-}\{\text{RACING}\}-\text{FINAL-WRITE} \), \( \text{ACT-}\{\text{GUARDED}\}-\text{WRITE} \), \( \text{ACT-}\{\text{RACING}\}-\text{FINAL-WRITE} \), \( \text{ACT-}\{\text{GUARDED}\}-\text{WRITE} \), and \( \text{ACT-}\{\text{RACING}\}-\text{FINAL-GUARDED-SPECULATION} \) the result is immediate since \( E = E' \).

\( \text{ACT-}\{\text{GUARDED}\}-\text{SPECULATION} \) are almost equally trivial since they simply add to \( E \), which clearly is a sensible extension. Letting \( E = F, x; \text{lock } t \) and \( E' = F, t; \text{effects}(s, E') \) gives the desired result for \( \text{ACT-RELEASE} \).

**Lemma 14 (Transitivity of sensibly extended environments).** Suppose \( E \vdash \sigma \vdash E' \). \( E' \) sensibly extends \( E \) and \( E'' \) sensibly extends \( E' \). Then, \( E'' \) sensibly extends \( E \).

**Proof.** Trivial from the definition of sensibly extended environments.

**Lemma 15 (Sensitive environment extensions preserve typability).** If \( E' \) sensibly extends \( E \) and \( E \vdash M \), then \( E' \vdash M \).

**Proof.** This follows from the definition of sensitive extensions and from Lemma 7.

Note, the previous lemma doesn’t hold for processes since removal of locks can invalidate if for processes.

**D.3 Decomposition/Composition**

**Lemma 16 (Context decomposition).** Suppose \( E \vdash C; \alpha \) and \( C \) is not a shape trap. Then \( \exists E' \) such that \( E \vdash C \vdash E' \) and \( E' \vdash \sigma \vdash E' \). Further, \( E' \) is a sensible extension of \( E \).

**Proof.** We prove this by induction on the context. First, let us note that as \( C \) is not a shape trap, we know by the definition of shape trap that the context \( C \) derived from \( C \) in the proof below is also not a shape trap.
Case $C = \square$: Immediate.

Case $C = C \land B$: By assumption and inversion, $C \land A \implies B$ must be typeable by CONTEXT-PAR, and therefore $(1) \implies B$ and $E \implies C \land A$. By induction, there exists $E'$ such that $(2) \implies C \land E' \implies B'$, $(3) E' \implies \neg A$, and $(4) E' \implies \neg A$ is a sensible extension of $E$. The result holds by $(3)$, $(4)$, and applying CONTEXT-PAR-1 to $(1)$ and $(2)$.

Case $C = \exists B$: Similar.

Case $C = (\nu \nu)C'$: By assumption and inversion, $(\nu \nu)C'[A]$ must be typeable by PROC-NU, giving $(1) E \vdash A$ and then $(2) E \vdash \neg A$ and $(3) E \vdash C' \land A$. By induction, there exists $E'$ such that $(2) E \vdash \neg A$ and $(3) E \vdash C' \land A$. By induction on $(3)$, there exists $(4) E' \implies C \land A$. By induction on $(2)$, we get $(5) E' \implies \neg A$. The result holds by $(3)$, $(4)$, and applying CONTEXT-PAR-2 instead of CONTEXT-PAR-1.

Case $C = \alpha C'$: By applying inversion and CONTEXT-ACTION to $(a)$, we get $(1) E \vdash \alpha \implies E$ and $(2) E \vdash \neg A$. We know by assumption that either (b') or (c) holds. In either case, we can apply the inductive hypothesis to $(2)$ since either $C'$ is a shape trap or $E \vdash A$. This gives $(3) E \vdash C'[A]$. Applying PROC-ACTION to $(1)$, we get $(4) E \vdash \alpha C'[A]$, the desired result.

Case $C = \exists B \implies \phi \implies B$: There are two cases for which typing rule will type $(a)$.

Case CONTEXT-SPECULATION-INITIAL: By applying inversion and CONTEXT-SPECULATION-INITIAL to $(a)$, we get $(1) E \vdash \alpha \implies E$ and $(2) E \vdash \neg A$. We know by assumption that either (b') or (c) holds. In either case, we can apply the inductive hypothesis to $(2)$ since either $C'$ is a shape trap or $E \vdash A$. This gives $(3) E \vdash C'[A]$. Applying PROC-SPECULATION to $(1)$, $(2)$, and $(3)$ gives $E \vdash \exists B \implies \phi \implies B$, as desired.

Case CONTEXT-SPECULATION-INITIAL-SHAPE-ERROR: By applying inversion and CONTEXT-SPECULATION-INITIAL-SHAPE-ERROR to $(a)$, we get $(1) E \vdash \alpha \implies E$ and $(2) E \vdash \neg A$. We know by assumption that either (b') or (c) holds. In either case, we can apply the inductive hypothesis to $(2)$ since either $C'$ is a shape trap or $E \vdash A$. This gives $(3) E \vdash C'[A]$. Applying PROC-SPECULATION-ERROR to $(1)$, $(2)$, $(3)$, and $(4)$ gives $E \vdash \exists B \implies \phi \implies B$, as desired.

Since CONTEXT-SPECULATION-INITIAL and CONTEXT-SPECULATION-INITIAL-SHAPE-ERROR are the only two judgments which could type $(a)$ under the current assumption of what shape $C$ has and the lemma holds for both of these judgments, then the lemma holds for this case.

Case $C = \exists B \implies \phi \implies C'$: There are two cases for which typing rule will type $(a)$.

Case CONTEXT-SPECULATION-FINAL: By applying inversion and CONTEXT-SPECULATION-FINAL to $(a)$, we get $(1) E \vdash B$, $(2) E \vdash \phi \implies E'$, and $(3) E \vdash \neg A$. By induction on $(2)$, we get $(4) E \vdash \phi \implies E$ and $(5) E \vdash B$. The result holds by (3), $(4)$, and applying CONTEXT-SPECULATION-FINAL to $(1)$, $(2)$, and $(3)$, we get the desired result.

Case CONTEXT-SPECULATION-ERROR: By applying inversion and CONTEXT-SPECULATION-ERROR to $(a)$, we get $(1) s \in \text{dom}(E)$, $(2) E \vdash p; D$, $(3) E \vdash C \land A$. We know by assumption that either (b') or (c) holds. In either case, we can apply the inductive hypothesis to $(3)$ since either $C'$ is a shape trap or $E \vdash A$. This gives $(4) E \vdash C'[A]$. Applying PROC-SPECULATION to $(1)$, $(2)$, and $(3)$ gives $E \vdash \exists B \implies \phi \implies C'[A]$, as desired.

Since CONTEXT-SPECULATION-FINAL and CONTEXT-SPECULATION-ERROR are the only two judgments which could type $(a)$ under the current assumption of what shape $C$ has and the lemma holds for both of these judgments, then the lemma holds for this case.

As we have proved all cases, we have proved the lemma.

D.4 Happens-before

In non-trap processes, the type system exactly captures the happens-before relation.

Lemma 18 (Happens-before, $s \implies C \implies E$).

(a) If $(t \downarrow p) <_{\text{cyc}} (t \downarrow p) \downarrow \text{then} E \vdash \exists (s; p)[p]$.  
(b) If $C$ is not a trap and $E \vdash \exists (s; p)[p]$ then $(t \downarrow p) <_{\text{cyc}} (t \downarrow p)$.  

Proof: By induction on $C$. The interesting cases are for speculation, and these follow from the definitions of trap and of justified speculation.

Lemma 19 (Justified read, $s \implies C \implies E$).

(a) If $(d \downarrow f) = \varnothing \downarrow \text{then} E \vdash \exists (f; C)[p]$ is a shape trap or $(g1) E \vdash \exists (f; C)[p] \vdash \varnothing T$.  
(b) If $(d \downarrow f) = \varnothing \downarrow \text{then} E \vdash \exists (f; C)[p] \vdash \varnothing T$.  

Proof: By induction on $C$. The interesting cases are for speculation, and these follow from the definitions of trap and of justified speculation.
Lemma 24 (Substitution). 1. If \( E; x : D \vdash M : C \) and \( E \vdash v : D \), then \( E \vdash M[[v]] : C \).

2. Assume \( E \vdash v : c.D, \) \( \text{mtype}(M.d) = c.\tilde{s}.S \rightarrow b.S, E \vdash \tilde{v} : \tilde{s}.S \rightarrow b.S, E, x : b.S \vdash M ; a.T, E \vdash W : b.S, \) and \( W \) is the result of \( p.m(\tilde{v}). \) Then \( E \vdash M[[\tilde{v}]] : a.T. \)

3. Assume \( E \vdash v : c.D, \) \( \text{mtype}(M.d) = c.\tilde{s}.S \rightarrow b.S, mbody(M.d) = \lambda \tilde{y}.M, \) and \( E \vdash \tilde{v} : \tilde{s}.S \). Then \( E \vdash M[[\tilde{v}]] : a.T. \)

Each of these can be proved using induction on \( E \vdash M : C \).

We now explicitly state the rules for defining \( A \geq B \) as the least precongruence on processes that satisfy the axioms in Equation 6.1. They are as follows:

- (s-transitivity): \( A \geq A' \) \( A' \geq A'' \) \( \Rightarrow A \geq A'' \)
- (s-ctx-action): \( A \geq A' \) \( A \geq A'' \) \( \Rightarrow A' \geq A'' \)
- (s-ctx-speculation): \( A \geq B \) \( A \geq B' \) \( \Rightarrow A \geq B' \)
- (s-ctx-constant): \( A \geq A' \) \( \Rightarrow A \geq A' \)
- (s-ctx-pass): \( A \geq A' \) \( \Rightarrow A \geq A' \)
- (s-ctx-predicate): \( A \geq A' \) \( \Rightarrow A \geq A' \)
- (s-ctx-speculation): \( A \geq A' \) \( \Rightarrow A \geq A' \)

Proposition 25 (Structural order preserves typing). \( E \vdash A, \) and \( A \geq B, \) then \( E \vdash B. \)

Proof. Assume that \( E \vdash A, \) and \( A \geq B. \) We prove this by induction on the structural order rules found in Equation 6.1 and on the ones explicitly stated above. For each case, we then use induction on the typing rules which are valid for that structural order.

Case 5-s-reflexivity: Assume that \( A = B. \) Therefore, the results holds trivially.

Case 5-s-transitivity: By induction on \( A \geq B, \) we get (1) \( A \geq A' \) and (2) \( A' \geq B. \) By the induction hypothesis on (1) and since \( E \vdash A, \) (3) \( E \vdash A'. \) Now, this together with the induction hypothesis on (2), means that (4) \( E \vdash B, \) as desired.

Case 5-s-ctx: In each case, the result follows using induction and induction.

Case 5-s-\( \text{nu-free} (\rightarrow) \): Assume that \( B = A(p)(q)(\tilde{p} : (\tilde{q} : \text{true}) p). \) We can further safely assume that \( \alpha \) \( \tilde{q} \) is defined, \( \beta \) \( \tilde{p} \) \( \text{ dom}(E), \) and that \( \gamma ) \) \( \tilde{q} \) \( \text{free} \) \( \text{ variables}. \) Otherwise, we can simply pick a different \( p \) and \( \delta \) which do satisfy these assumptions. We know by assumption that \( E \vdash A, \) so (1) \( E \vdash \alpha. \) By the rules for typing \( E \vdash \alpha \) and \( E \vdash \beta, \) we know that (2) \( E \vdash \gamma. \) By (a), (2), and (c), we know that (3) \( E \vdash (\tilde{q} : \text{free}). \) As we have (1), (3), and (b), we know that (4) \( E, p : (\tilde{q} : \text{free}). \) \text{returns} \( \text{Unit} \vdash \alpha. \) Further, by applying \( \text{proc-free} \) to (1) and (4), we have (5) \( E, p : (\tilde{q} : \text{free}). \) \text{returns} \( \text{Unit} \vdash \gamma. \) We can also apply \( \text{proc-to} \) (1) and (4), giving us (6) \( E, p : (\tilde{q} : \text{free}). \) \text{returns} \( \text{Unit} \vdash \gamma. \) By applying \( \text{proc-to} \) (7), we get (8) \( E \vdash (\tilde{v} : (\tilde{q} : \text{free}) : \text{true} p). \) Finally, by applying \( \text{proc-to} \) \( E \vdash \alpha \) and (8), we get (9) \( E \vdash A(p)(q)(\tilde{p} : (\tilde{q} : \text{true}) p). \) as desired.

Case 5-s-\( \text{nu-free} (\rightarrow) \): Assume that \( A = A'(p)(q)(\tilde{p} : (\tilde{q} : \text{false}) p). \) We can further safely assume that \( A = A'(p)(q)(\tilde{p} : (\tilde{q} : \text{true}) p). \) We can further safely assume that \( A = A'(p)(q)(\tilde{p} : (\tilde{q} : \text{true}) p). \) We can further safely assume that \( A = A'(p)(q)(\tilde{p} : (\tilde{q} : \text{true}) p). \) We can further safely assume that \( A = A'(p)(q)(\tilde{p} : (\tilde{q} : \text{true}) p). \) We can further safely assume that \( A = A'(p)(q)(\tilde{p} : (\tilde{q} : \text{true}) p). \)
Case S-PAR-PAR ($\Rightarrow$):

Case S-PAR-PAR ($\Rightarrow$):

Case S-PAR ($\Rightarrow$):

Case S-PAR ($\Rightarrow$):

Trivial through application of inversion and PROC-PAR.

Case S-NU-NU ($\Rightarrow$):

Trivial through application of inversion and PROC-NU, Lemma 7, Corollary 8, Lemma 9, and Corollary 10.

Case S-NU-PREFIX ($\Rightarrow$):

Let $A = (\nu p)\alpha A$ and $B = p(\nu p)\alpha A'$ such that $p \notin \text{fn}(\alpha)$. Applying inversion and PROC-NU to $E \vdash A$ gives (1) $E, \nu p \vdash p \in \text{dom}(\alpha)$. Applying inversion and PROC-NU to (1) gives (2) $E, \nu p \vdash \alpha A'$. Applying Corollary 8 followed by Corollary 10 to (2) gives (3) $E \vdash \alpha A'$. Applying PROC-NAC on (4) and (7) gives $E \vdash \alpha (\nu p)\alpha A'$, the desired result.

Case S-NU-PREFIX ($\Rightarrow$):

Similar.

Case S-PREFIX-PAR:

Let $A = A'((\alpha A')A)$ and $B = \alpha A'(A'\alpha A')$. We know by applying inversion and PROC-PAR to (1) that $E \vdash A'$ and (2) $E \vdash \alpha A'$ hold. Applying inversion and PROC-NAC to (2) gives (3) $E \vdash \alpha A'. E \vdash A'$. We know by the well-formedness of $A$ that $\alpha$ is contiguous. Thus we know that if $\alpha$ is a synchronization action, then it can’t overlap with $A$. Thus, by Corollary 8, we have (5) $E \vdash \alpha A'$. Applying inversion and PROC-PAR to (5) gives (6) $E \vdash A'$ and (7) $E \vdash \alpha A'$. Applying PROC-PAR to (6) and (4) gives (8) $E \vdash \alpha A'^3$. Applying PROC-NAC on (3) and (8) gives (9) $E \vdash \alpha (A'\alpha A')$, as desired.

Case S-PREFIX:

Let $A = \alpha A'$ and $B = \alpha A'$ such that $\alpha = \beta$. We know by assumption that $E \vdash A'$. By successive applications of inversion and PROC-NAC, we have (1) $E \vdash \alpha A'$ and (2) $E \vdash A'$. By applying Lemma 22, we have (3) $E \vdash \beta A'$, as desired.

Case S-PREFIX-SPECULATION:

Let $A = (\alpha A')A$ and $B = \alpha A'(A'\alpha A')$ such that (1) $\phi \alpha A' \phi = \alpha A' \phi = A'$, (2) $\phi \alpha A' = \alpha A' \phi = A'$, and (3) $\alpha \phi A' = \alpha A' \phi = A'$. As $E \vdash A'$, there are two typing rules which can be applied here, PROC-SPECULATION and PROC-SPECULATION-SHAPE-ERROR. We consider both cases to complete the proof of this case.

Case PROC-SPECULATION:

Assume that PROC-SPECULATION types $E \vdash A$. By inversion and PROC-SPECULATION, we have (3a) $E \vdash \alpha A'$, (4a) $E \vdash \beta E'$, and (5a) $E \vdash \alpha A'$. Applying inversion and PROC-NAC to (2a) gives (6a) $E \vdash \alpha A'$ and (7a) $E \vdash A'$. Applying PROC-NAC to (5a) gives (8a) $E \vdash \alpha A'$ and (9a) $E \vdash A'$. Applying Lemma 23 to (1), (2), (4a), (8a), and (6a) gives (11a) $E \vdash A' \phi E' = E' \phi A'$.

Applying PROC-NAC to (7a), (11a), and (7a) gives (12a) $E \vdash \alpha A' \phi \alpha A'$. Applying PROC-NAC (6a) and (12a) gives $E \vdash \alpha A' \phi \alpha A'$, as desired.

Case PROC-SPECULATION-SHAPE-ERROR:

Assume that PROC-SPECULATION-SHAPE-ERROR types $E \vdash A$. Thus, $\phi = (\alpha \beta \gamma)$. By inversion and PROC-SPECULATION-SHAPE-ERROR, we have (3b) $s \in \text{dom}(E)$, (4b) $E \vdash \beta \gamma E'$, (5b) $E \vdash \beta \gamma E'$, (6b) $E \vdash \alpha A'$. Applying inversion and PROC-NAC to (6b) gives (7b) $E \vdash \alpha A'$. It is clear that if (3b), (4b), and (5b) hold, then $E \vdash \alpha A'$. This follows from the definition of a shape trap, $E \vdash \alpha A'$.

Since we showed that for both cases of typing $E \vdash A$ results in $E \vdash B$ being typed, we have shown that this case is satisfied as well.

Case S-NU-NU-PREFIX ($\Rightarrow$):

Let $A = (\nu p)(\exists A' \phi A' \phi A' \phi A')$ and $B = (\exists A' \phi A' \phi A' \phi A')$ such that $p \notin \text{fn}(\phi)$. Applying inversion and PROC-NAC to $E \vdash A$ gives (1) $E, \nu p \vdash p \in \text{dom}(\phi)$. Applying inversion and PROC-NAC to (1) gives (2) $E, \nu p \vdash \phi A'$. Applying Corollary 8 followed by Corollary 10 to (2) gives (3) $E \vdash \phi A'$. Applying PROC-NAC on (4) and (7) gives $E \vdash \phi (\nu p)A'$, the desired result.

D.7 Reduction preserves typing

Lemma 26 (Shape traps preserve typability). If $(1) \theta \vdash C \subseteq [M]$, $(2) C$ is a shape trap, and $(3) \subseteq C \subseteq [M]$, then $\theta \vdash C \subseteq [M]$.

Proof. Let us assume (1), (2), and (3). Then we know that $C'$ is also a shape trap since traps can never be removed. So, by the definition of a shape trap, $\theta \vdash \exists D \vdash E$ and $C' \subseteq \exists D \vdash E$. Applying Lemma 23 to (1), (2), (4a), and (6a) gives (12a) $E \vdash A' \phi E' = E' \phi A'$. Applying PROC-NAC to (6a) and (12a) gives $E \vdash \alpha A' \phi \alpha A'$, as desired.

Lemma 27 (Thread reduction introduces a shape trap or sensibly extends the environment). If $\theta \vdash C \subseteq [M]$, $\theta \vdash C' \subseteq [M]$, and $\theta \vdash C \subseteq E$, then either (1) $\theta \vdash C \subseteq E$ or (2) $\theta \vdash C' \subseteq E$.

Proof. If $C$ is a shape trap, we are done since $C'$ must also be a shape trap since traps can never be removed. If a shape trap is introduced to $C'$ during the reduction step, we are again done. Therefore the interesting part of the proof is when $C$ and $C'$ are not shape traps. This can be proven by induction on reduction rules found in Equation 6.1, Equation 6.1, and Figure 6.

Lemma 28 (Thread return type existence). If $E \vdash s[M]$, then $E \vdash E'. s \vdash \text{returns} A t$, for some $E'$. 

Proof. Assume \( E \vdash s \langle M \rangle \). Applying inversion and PROC-THREAD gives (1) \( E \triangleright s \text{ returns } a \langle T \rangle \) and (2) \( E \triangleright v \in M : a \rangle. Since we have (1) we can reorder \( E \) so that \( E \vdash E', s \text{ returns } a \langle T \rangle \), as desired.

\[ \square \]

Lemma 29 (Thread return types). If \((v_1) \in C\), then if \( E' \vdash C' \rightarrow E', s \text{ returns } a \langle T \rangle \), then \( E \vdash C > C' \rightarrow E', s \text{ returns } b \langle S \rangle \). If \((v_2) \notin C\), then if \( E, s \text{ returns } a \langle T \rangle \), then \( E \vdash E', s \text{ returns } b \langle S \rangle \).

Proof. The first case follows from CONTEXT-NU. The second case follows from the fact that \( C \) doesn’t modify the return type of a thread.

\[ \square \]

We now state and prove preservation, also known as subject reduction.

Proposition 30 (Preservation). If \( \emptyset \vdash A \) and \( A \rightarrow B \), then \( \emptyset \vdash B \).

Proof. Assume that \((1) \emptyset \vdash A \) and \((2) A \rightarrow B \). We prove this by induction on the reduction rules found in Equation 6.1, Equation 6.1, and Figure 6. That is, we prove this by induction on (2). We start with the case of R-STRUCTURAL-ORDER.

Case R-STRUCTURAL-ORDER: Assume that R-STRUCTURAL-ORDER satisfies (2). Then by assumption, we know that \((3) A \rightarrow A', (4) A' \rightarrow B' \), and (5) \( B' \geq B \). Applying Proposition 25 on (1) and (3) gives (6) \( \emptyset \vdash A' \). Applying the inductive hypothesis on (6) and (4) gives (7) \( \emptyset \vdash B' \). Applying Proposition 25 on (7) and (5) gives (8) \( \emptyset \vdash B \), as desired.

We partition the remaining cases into two sets that we need to consider. One set of cases is comprised of the reduction rules which modify the context while reducing, namely R-METHOD-CONTEXT, and R-LOCK-CONTEXT. The other set consists of all remaining cases and do not modify the context. Let us first consider the cases for the rules which modify the context, as they are more interesting.

Case R-METHOD-CONTEXT: Assume that R-METHOD-CONTEXT satisfies (2). So, \((3) A \vdash C \vdash \{s | v \in \langle N \rangle, M \}\), \((4) B \vdash C \vdash \{s | v \in \langle N \rangle, M \}\). By the equalities of (3) and (4) combined with (2), we have (5) \( C[\{s | v \in \langle N \rangle, M \}] \rightarrow C[\{s | v \in \langle N \rangle, M \}] \), and (6) \( C[\{s | v \in \langle N \rangle, M \}] \rightarrow C[\{s | v \in \langle N \rangle, M \}] \). Now there are two cases to consider. Either \( C \) is a shape trap or it isn’t. If it is, then by Lemma 26, \( C \vdash C \vdash \{s | v \in \langle N \rangle, M \} \). Now since \( B \vdash C \vdash \{s | v \in \langle N \rangle, M \} \), we have \( B \vdash \emptyset \), as desired. Thus, if \( C \) is a shape trap, we are done. We proceed assuming that \( C \) is not a shape trap.

Applying Lemma 16 on (1) and (3), we have \((7) \emptyset \vdash E \triangleright v \in \langle N \rangle \). Applying inversion and PROC-THREAD on (8), we get (10) \( E \vdash s \text{ returns } a \langle T \rangle 

Case R-LOCK-CONTEXT: Similar. The only differences are the use of STAT-LOCK-CONTEXT instead of STAT-METHOD-CONTEXT and the handling of the additional shape trap from the judgment, namely \( E \vdash \text{s returns } a \langle T \rangle \rightarrow \square \text{Lock} \). Of course, this can be handled in the same way that (14) \( E \vdash s \text{ returns } a \langle T \rangle \), \( x : b > S \rightarrow M : a \rangle \) from the previous case is handled.

We now have to prove the proposition is true for the remaining cases. Assume that one of the remaining reduction rules satisfies (2). Let us assume for this case that \((3) A \vdash C \vdash \{s | v \in \langle N \rangle, M \}\). By inspecting the rule for this case, we know that \( C \vdash C' \). Thus, we have (5) \( B \vdash C'[\{s | v \in \langle N \rangle, M \}] \). Now there are two cases to consider. Either \( C \) is a shape trap or it isn’t. If it is, then by the fact that Lemma 26 holds, \( \emptyset \vdash C'[\{s | v \in \langle N \rangle, M \}] \). Now since (5) \( B \vdash C'[\{s | v \in \langle N \rangle, M \}] \), we have \( \emptyset \vdash B \), as desired. Since \( C \) is a shape trap, we have \( \emptyset \vdash B \), as desired. Since we arbitrarily chose the reduction rule from the remaining ones, we know that this is the case for all of them. So for each of the remaining cases, we proceed assuming that \( C \) is not a shape trap. We consider each case in turn beginning with the most interesting ones.

Case R-SPECULATION-CLOSE: Assume that R-SPECULATION-CLOSE satisfies (2). Let us assume \((3) A \vdash C \vdash \{s | v \in \langle N \rangle, M \}\) and (4) \( B \vdash C'[\{s | v \in \langle N \rangle, M \}] \). By the equalities of (3) and (4) combined with (2), we have (5) \( C[\{s | v \in \langle N \rangle, M \}] \rightarrow C[\{s | v \in \langle N \rangle, M \}] \), and (6) \( C[\{s | v \in \langle N \rangle, M \}] \rightarrow C[\{s | v \in \langle N \rangle, M \}] \). Applying Lemma 16 to (1), we know that \((7) \emptyset \vdash C \vdash E \). Applying Lemma 29 to (16) gives (17a) \( E \vdash \{s | v \in \langle N \rangle, M \} \). By inversion and PROC-SPECULATION on (8), we have (9) \( E \vdash \{s | v \in \langle N \rangle, M \} \). We note that PROC-SPECULATION-SHAPE-ERROR cannot be applied here since (6) holds. Applying Lemma 20 to (6), (10), and (11) \( E \vdash B \). Applying Lemma 17 to (12) and (11) \( E \vdash \{s | v \in \langle N \rangle, M \} \). By the equality of (4), we know \( (14) \emptyset \vdash B \), as desired.

Case R-OPENING-PHENOMENA: Assume that R-OPENING-PHENOMENA satisfies (2). Let us assume \((3) A \vdash C \vdash \{s | v \in \langle N \rangle, M \}\) and (4) \( B \vdash C'[\{s | v \in \langle N \rangle, M \}] \). We now type the synchronization actions. By applying inversion on PROC-RUNNABLE and PROC-FREE to (13) and (14), we have \((17) E \vdash \{s | v \in \langle N \rangle, M \} \). Applying Lemma 16, inversion and PROC-THREAD on (13) and (14), we have (18) \( E \vdash \{s | v \in \langle N \rangle, M \} \). Applying ACT-RELEASE, and ACT-ACQUIRE again on (14), (18), and (19), we get (20) \( E \vdash \{s | v \in \langle N \rangle, M \} \).
\[E = F, s:Z \vdash \text{effects}(t, E),\]
\[Z = \{\xi \mid E \triangleright \text{onacq } \xi \quad \text{and} \quad E \triangleright \xi \} ,\]
\[F' = F', t: \text{effects}(s, F),\]
\[E' = E', p: \text{lock } t: \text{effects}(t, E'),\]
\[Z' = \{\xi' \mid E' \triangleright \text{onacq } \xi' \quad \text{and} \quad F' \triangleright \xi'\} .\]

We now type \(p[N'_2/\text{thread}]\) and \(s[M'_1/\text{thread}]\). By the definitions of \(F', F'', (6), (18), (19), E'' \vdash \triangleright \text{D} \). Noting that the type of \(\text{mtype}(D, \text{run}) = \triangleright \text{D} \rightarrow \triangleright \text{Unit}\), we can apply Lemma 24 on (7) followed by \(\text{proc-thread} \rightarrow E'' \vdash p[N'_2/\text{thread}]\). Starting from (15), applying inversion and \(\text{stat-method}\), Lemma 24, and \(\text{proc-thread}\) gives (24) \(E' \vdash s[M'_1/\text{thread}]\). We can combine the processes and actions as follows. Starting with (23), we can apply \(\text{proc-action}\) on (22), (21), and (20), followed by \(\text{proc-pass}\) on (24). This gives (25) \(E' \vdash s[M'_1/\text{thread}]\). By the equalities of (3) and (4) combined with (2), we have (5) \(\text{free } s\vdash \text{val } x = \text{new } \text{D}(\triangleright \text{D})\). By the equalities of (4) to (27), we have (28) \(\triangleright \text{B} = \triangleright \text{B} = \text{as desired.}\)

Case \(\text{R-NEW}\): Assume that \(\text{R-NEW}\) satisfies (2). Let us assume (3) \(A = C[A']\), (4) \(B = C[B']\), where \(A' = \text{free } s\vdash \text{val } x = \text{new } \text{D}(\triangleright \text{D})\) and \(B' = \text{runnable } p \vdash \text{get}(p, \text{null}) = x\) \(\vdash \text{null}\). By the equalities of (3) and (4) combined with (2), we have (5) \(\text{free } s\vdash \text{val } x = \text{new } \text{D}(\triangleright \text{D})\). Further, by applying the lemmas, we have (6) \(\triangleright \text{C} \vdash p: \triangleright \text{D}\). (7) \(\text{final}(\text{D}) = \text{f}, \text{nonfinal}(\text{D}) = \overline{\text{f}}\), and (9) \(\text{D} \). By the equalities of (4) to (27), we have (28) \(\triangleright \text{B} \rightarrow \text{as desired.}\)

Case \(\text{R-FIELD-WRITE}\): Assume that \(\text{R-FIELD-WRITE}\) satisfies (2). This gives us (3) \(A = C[A'], (4) B = C[B']\), where \(A' = \text{free } s\vdash \text{val } x = \text{new } \text{D}(\triangleright \text{D})\). By the equalities of (3) and (4) combined with (2), we have (5) \(\text{free } s\vdash \text{val } x = \text{new } \text{D}(\triangleright \text{D})\). Further, by applying the lemmas, we have (6) \(\triangleright \text{C} \vdash p: \triangleright \text{D}\). (7) \(\text{final}(\text{D}) = \text{f}, \text{nonfinal}(\text{D}) = \overline{\text{f}}\), and (9) \(\text{D} \). By the equalities of (4) to (27), we have (28) \(\triangleright \text{B} \rightarrow \text{as desired.}\)

Case \(\text{STAT-RACING-WRITE}\): Assume that \(\text{STAT-RACING-WRITE}\) types (8). By inversion and \(\text{stat-racing}\), we know (9a) \(E \vdash t: \text{par} \vdash p: \text{D}, (10a) \text{free } (D) = \text{D} \triangleright \text{S}\). By the equalities of (3) and (4) combined with (2), we have (5) \(\text{free } s\vdash \text{val } x = \text{new } \text{D}(\triangleright \text{D})\). Further, by applying the lemmas, we have (6) \(\triangleright \text{C} \vdash p: \text{D}\). (7) \(\text{final}(\text{D}) = \text{f}, \text{nonfinal}(\text{D}) = \overline{\text{f}}\), and (9) \(\text{D} \). By the equalities of (4) to (27), we have (28) \(\triangleright \text{B} \rightarrow \text{as desired.}\)
Case $\text{R-CAST:}$ Assume that $R\text{-CAST}$ satisfies (2). Let us assume (3) $A = C[A']$. (4) $B = C[B']$, where $A' = s[\\text{val} x = \langle D|p; M \rangle]$ and $B' = s[M]$. By the equalities of (3) and (4) combined with (2), we have (5) $C[s[\\text{val} x = \langle D|p; M \rangle] \rightarrow C[s[M]]$, and (6) $C \ni p : D$. By applying Lemma 16, inversion, and PROC-THREAD on (1) and (3), we have (7) $\not\vdash C \supset E$, $E$ is a sensible extension of $\not\vdash$ (8) $E \ni s \text{returns } aT$, and (9) $E \vdash \langle D|p; M : aT \rangle$. Applying inversion and STAT-CAST gives (10) $E \vdash p : b\text{Object}$, and (11) $E \vdash bD \vdash M : aT$. By inversion and VAL-SUB-TYPE, (6), and (10) gives (12) $E \vdash p : bD$. Applying Lemma 24 on (11) and (12) followed by PROC-THREAD gives (13) $E \vdash M \langle \ell/\ell \rangle : aT$ and (14) $E \vdash s[M]$. By applying Lemma 17 on (7) and (14) and using the equality of (4), we have (15) $\not\vdash B$, as desired.

Case $\text{R-METHOD-CALL:}$ Assume that $R\text{-METHOD-CALL}$ satisfies (2). Let us assume (3) $A = C[s[\\text{val} x = p.m(\ell)]; M]$, (4) $B = C[s[\\text{val} x = \langle N[\ell/\ell\langle\ell\rangle]; \ell\rangle M]]$. By the equalities of (3) and (4) combined with (2), we have (5) $C[s[\\text{val} x = p.m(\ell)]; M] \rightarrow C[s[\\text{val} x = \langle N[\ell/\ell\langle\ell\rangle]; \ell\rangle M]]$. By the assumptions for the rule, we have (6) $C \ni p : D$, (7) $\text{mbody}(D.m) = \lambda x. N$, (8) $D$ not reserved, and (9) $m \neq \text{start}$. By applying Lemma 16, and inversion and PROC-THREAD on (1) and (3), we have (10) $\not\vdash C \supset E$, $E$ is a sensible extension of $\not\vdash$ (11) $E \ni s \text{returns } aT$ and (12) $E \vdash s[\\text{val} x = p.m(\ell); M] : aT$. Applying inversion and STAT-METHOD and then Lemma 24 on (12) gives (13) $E \vdash x : bS \vdash \langle N[\ell/\ell\langle\ell\rangle]; \ell\rangle M$. Reordering the environments in (13) and (14), applying STAT-METHOD-CONTEXT and PROC-THREAD gives (15) $E \vdash s[\\text{val} x = \langle N[\ell/\ell\langle\ell\rangle]; \ell\rangle M]$. By applying Lemma 17 on (10) and (15) and using the equality of (4), we have (16) $\not\vdash B$, as desired.

Case $\text{R-OPERATOR:}$ Assume that $R\text{-OPERATOR}$ satisfies (2). Then we have $C[s[\\text{val} x = \ell; M]] \rightarrow C[s[M]]$. This case can be handled similarly to the other cases. Apply Lemma 16, inversion and PROC-THREAD, inversion and STAT-METHOD, inversion and STAT-RETURN, Lemma 24, PROC-THREAD, and then Lemma 17 to get the desired result.

Case $\text{R-FIELD-READ:}$ Assume that $R\text{-FIELD-READ}$ satisfies (2). Then we have $C[s[\\text{val} x = \ell; M]] \rightarrow C[s[M]]$. This case can be handled similarly to $R\text{-METHOD-CALL}$. Apply Lemma 16, inversion and PROC-THREAD, inversion and STAT-OPERATOR, Lemma 24, PROC-THREAD, and then Lemma 17 to get the desired result.

Since we showed that all reduction rules satisfy the proposition, we know that subject reduction holds. $\square$