A Model of Transactional Programming

Kousha Etessami†  Patrice Godefroid‡

1 University of Edinburgh, kousha@inf.ed.ac.uk
2 Microsoft Research, pg@microsoft.com

Abstract. There has been a lot of recent research on transaction-based concurrent programming, aimed at offering an easier concurrent programming paradigm that enables programmers to better exploit the parallelism of modern multi-processor machines, such as multi-core microprocessors. We introduce Transactional State Machines (TSMs) as an abstract finite-data model of transactional shared-memory concurrent programs. TSMs are a variant of concurrent boolean programs (or concurrent extended recursive state machines) augmented with additional constructs for specifying potentially nested transactions. Namely, some procedures (or code segments) can be marked as transactions and are meant to be executed “atomically”, and there are also explicit commit and abort operations for transactions. The TSM model is non-blocking and allows interleaved executions where multiple processes can simultaneously be executing inside transactions. It also allows nested transactions, transactions which may never terminate, and transactions which may be aborted explicitly, or aborted automatically by the run-time environment due to memory conflicts.

We show that concurrent executions of TSMs satisfy a correctness criterion closely related to serializability, which we call stutter-serializability, with respect to shared memory. We initiate a study of model checking problems for TSMs. Model checking arbitrary TSMs is easily seen to be undecidable, but we show it is decidable in the following case: when recursion is exclusively used inside transactions in all (but one) of the processes, we show that model checking such TSMs against all stutter-invariant ω-regular properties of shared memory is decidable.

1 Introduction

In recent years serious obstacles have impeded further increases to the clock speed of microprocessors. By contrast, the number of transistors on a chip has continued to grow. This has led to the advent of multi-core microprocessors, where multiple CPUs are placed on a single chip. If software is to take full advantage of the parallelism afforded by such multiprocessor machines, concurrent shared-memory programming will need to become far more prevalent. Unfortunately, concurrent programming is notoriously difficult and despite decades of research the available concurrent programming paradigms (e.g., using locks,
semaphores, monitors, etc.) have proven to be hard to use and highly error prone. We need easier paradigms if concurrent programming is to become practical and commonplace.

One paradigm on which a great deal of recent research has focused is transacational programming, which attempts to adapt powerful transaction processing technology, developed originally for database concurrency control, to the setting of general purpose concurrent programs. Roughly speaking, transactions are marked code segments that are to be executed “atomically”. The goal of such research is to use transactions as the main enabling construct for shared memory concurrent programming, replacing more conventional but low level constructs such as locks. High level transactional code could in principle then be compiled down to machine code for the shared memory machine, as long as the machine provides certain needed low level atomic operations (such as atomic compare-and-swap). Already, a number of languages and libraries for transactions have been implemented (see, e.g., [15] which surveys many implementations). Much of this work however lacks precise formal semantics specifying exactly what correctness guarantees are provided by the transactional framework. Indeed, there often appears to be a tension between providing strong formal correctness guarantees and providing an implementation flexible and efficient enough to be deemed useful. Even if transactional constructs were themselves given clear semantics, there would still remain the important task of verifying specific desired properties of specific transactional programs.

The aim of this paper is to provide a state-machine based formal model of transactional concurrent programs, and thus to facilitate an abstract framework for reasoning about them. In order for such a model to be useful, firstly, it must be close enough to existing transactional paradigms so that, in principle, such models could be derived from actual transactional programs via a process of abstraction akin to that for ordinary programs. Secondly, the model should be simple enough to enable (automated) reasoning about such programs.

So, what is a “transaction”? Syntactically, transactions are marked code segments, e.g., demarcated by “atomic {...}”, or, more generally, they are certain procedures which are marked as transactional. (Simple examples of transactional concurrent shared-memory programs are given in Figure 1. These examples will be discussed later.) But what is the semantics? An informal and naive view of their semantics is given by so-called “single-lock semantics” (see, e.g., [15]), which says that during concurrent execution each executed transaction should appear “as if” it is executing serially without any interleaving of the operations of that transaction with other transactions occurring on other processes. In other words, it should appear “as if” executing each transaction requires every process to acquire a single global transaction lock and to release that lock only when the transaction has completed. The problem with this informal semantics has to do with precisely what is meant by “as if”. A semantics which literally assumes that every concurrent execution proceeds via a single lock, literally rules out any interleaving of transactions on different processes. This is too restrictive and disallows a lot of useful concurrency. It also violates the intended
non-blocking nature of the transactional paradigm, and ignores other features, such as the fact that typically transactions can be aborted either explicitly or automatic by the run time system due to memory conflicts. After all, if single-lock semantics was literally intended, then it would not be difficult to implement, but would obviously be hopelessly inefficient. Of course, designers of transactional frameworks would object to this literal interpretation of “as if”. Single-lock semantics is not literally what is intended. Rather, a weaker semantics is intended, but phrasing a simple formal semantics which captures precisely what is desired and leaves sufficient flexibility for an efficient implementation is itself a non-trivial task. Standard correctness notions such as serializability, which are used in database concurrency control, are not directly applicable to this setting without some modification. This is because in full-fledged concurrent programming it is no longer the case that every operation on memory is done via a transaction consisting of a block of (necessarily terminating) straight-line code. The “transactional program” running on each process may consist of a mix of transactional and non-transactional code, transactions may be nested, and moreover some transactions (which are programs themselves) may never halt. When adapting correctness criteria to this setting, one needs to take careful account of all these subtle differences.

In this paper, we propose Transactional State Machines (TSMs) as an abstract finite-data model of transactional shared-memory concurrent programs. We define the concept of stutter-serializability, which we feel captures in a clean and simple way a desired correctness criterion, and we show that TSMs satisfy this condition. We also study model checking of TSMs.

The TSM model is non-blocking and allows interleaved executions where multiple processes can simultaneously be executing inside transactions. It also allows nested transactions, transactions which may never terminate, and transactions which may be aborted explicitly, or aborted automatically due to memory conflicts. Our definition of TSMs is based on two natural assumptions which are close in spirit to assumptions used in transactional memory systems:

1. We assume the availability of a (hardware or software implemented) atomic multi-word compare-and-swap operation, $\text{CAS}(\vec{x}, \vec{x}', \vec{y}, \vec{y}')$, which compares the contents of the vector of memory locations $\vec{x}$ to the contents of the vector of memory locations $\vec{x}'$, and if they are the same, it assigns the contents of the vector of memory locations $\vec{y}'$ to the vector of memory locations $\vec{y}$.\(^1\)

\(^1\) This compare-and-swap (CAS) operation is of course far more general than the conventional CAS operation offered by some hardware platforms, which typically only allows comparison of a single memory location to a given value, and if they are the same, assigns a new value to the same memory location. We shall need the extra power in order to establish strong correctness properties for TSMs. However, obviously, if such a strong atomic operation is “mis-used” it is unrealistic to expect it to be efficiently implementable. Indeed, this and other aspects of the TSM model suggest naturally that one should restrict the use of transactions in concurrent programs to “small” transactions that require only “small” CAS operations, in order to
2. We assume a form of strong isolation (strong atomicity).² Specifically, there must be minimal atomic operation units on all processes, such that these atomic units are indivisible in a concurrent execution, meaning that a concurrent execution must consist precisely of some interleaved execution of these atomic operations from each process. Thus “atomicity” of operations must hold at some level of granularity, however small. Without such an assumption, it is impossible to reason about asynchronous concurrent computation via an interleaving semantics, which is what we wish to do.

Based on these two assumptions, we can now give an informal description of the TSM model. TSMs are concurrent boolean programs with procedures, except that some procedure calls may be transactional (and such calls may also be nested arbitrarily). Transactional calls are treated differently at run time. After a transactional call is made, the first time any part of shared memory is used in that transaction, it is copied into a fixed local copy on the stack frame for that transaction. A separate, mutable, copy (valuation) of shared variables is also kept on the transactional stack frame. All read/write accesses (after the first use) of shared memory inside the transaction are made to the mutable copy on the stack, rather than to the universal copy. Each transaction keeps track (on its stack frame) of those shared memory variables that have been used or written during the execution of the transaction. Finally, if and when the transaction terminates, we use an atomic compare-and-swap operation to check that the current values in (the used part of) the universal copy of shared memory are exactly the same as their fixed copy on the stack frame, and if so, we copy the new values of written parts of shared memory from the mutable copy on the stack frame to their universal copy. Otherwise, i.e., if the universal copy of used shared memory is inconsistent with the fixed copy for that transaction, we have detected a memory conflict and we abort that transaction.

The key point is this: if the compare-and-swap operation at the end of a transaction succeeds and the transaction is not aborted, then we can in fact “schedule” the entire activity of that transaction inside the “infinitesimal time slot” during which the atomic compare-and-swap operation was scheduled. In other words, there exists a serial schedule for non-aborted transactions, which does not interleave the operations of distinct non-aborted transactions with each other. This allows us to establish the stutter-serializability property for TSMs.

The above description is over-simplified because, e.g., TSMs also allow nested transactions and there are other technicalities, but it does describe some key aspects of the model. We describe the full model in detail in section 2 (and the appendix). We show that TSMs are stutter-serializable in section 3. We study model checking for TSMs in section 4, and show that, although model checking for general TSMs is undecidable, there is an interesting fragment for which it

² Terms like strong/weak isolation/atomicity (see [7, 15]) have in fact been used by different authors to mean subtly different things, and must be defined carefully with respect to the model.
Initially, $x = 0$

Process 1
atomic {
  $r_1 = x$;
  $x = 1$;
  $x = 2$;
}

Can $r_1 == 1$? No.

Initially, $x = y = 0$

Process 1
atomic {
  $r_1 = y$;
  $y = 1$;
  atomic {
    if ($x == 0$)
      $x = 1$;
    abort;
  }
}

Can $r_1 == 1$? No.

Initially, $x = y = 0$

Process 1
atomic {
  $r_1 = x$;
  $x = 1$;
  $r_2 = y$;
  $y = 1$;
}

Can $r_1 == 1$, $r_2 == 0$? No.

Fig. 1. Examples

remains decidable. Namely, when recursion is exclusively used inside transactions in all (but one) of the processes, we show that model checking such TSMs against all stutter-invariant $\omega$-regular properties of shared memory is decidable.

**Examples.** Figure 1 contains simple example transactional programs (adapted from [12]). Transactions are syntactically defined using the keyword `atomic`. With each example, we describe the possible effect, in our TSM model, on the variables $r_1$ (and $r_2$) at the end of the example’s execution. As mentioned, in the TSM model the execution of transactions on multiple processes can interleave, and moreover the execution of transactional and non-transactional code can also interleave. So, in the leftmost example, what happens if the non-transactional code executed by Process 2 executes before the transaction on Process 1 has completed? In the TSM model, Process 2 would read the value of the shared variable $x$ from a universal copy of shared memory which has not yet been touched by the executing transaction on Process 1. If Process 1 completes its transaction and commits successfully, then the final value 2 is written to this universal copy of $x$, and thereafter Process 2 could read this copy and thus it is possible that $r_1 == 2$ after this program has finished. We note that $r_1 == 1$ would be possible at the end under forms of weak atomicity, e.g., if `atomic` was implemented as a synchronized block in Java (see [12]). The middle example in Figure 1 contains an explicit abort. In the TSM model, all write operations on shared variables performed by a transaction only have an effect on (the universal copy of) shared memory if the transaction successfully commits. Otherwise they have no effect, and are not visible to anyone after the transaction has been aborted. This is a form of deferred update as opposed to direct update ([15]), where writes in an aborted transaction do take effect, but the abort overwrites them with the original values. In that case, such a write might be visible to non-transactional code and $r_1$ might have the value 1 at the end of execution of this example. Note that our semantics for TSMs does not take into account possible re-orderings that may be performed by standard compilers. For instance, compilers are usually allowed to re-order read operations, such as those performed by Process 2 in the rightmost example in Figure 1. These ordering issues [12] are not discussed in this paper. Indeed, extending TSMs to incorporate notions of reordering would complicate the model and detract from our main goal of
having a clean abstract reference model which brings to light the salient aspects
of transactional concurrent programs.

**Discussion of related work and TSMs.** There is an extensive literature on
Transactional Memory and there are already many prototype implementations
(see the online bibliography [8], and see the recent book by Larus and Rajwar [15]). Most of this work discusses how to implement transactional memory either
in hardware or software, from a systems point of view with the main emphasis
on performance. Only recently (e.g., see [17, 1]) some researchers have started to
formalize and study the semantics of transactional memory implementations, in
order to clarify subtle semantics distinctions between various implementations
and the interface between these implementations and higher-level “transactional
programs” running on top of them. A particular focus in [17, 1] is on finding
some meaningful semantics in the difficult setting of weak isolation/atomicity,
where implementations do not detect conflicting accesses to shared memory be-
tween non-transactional and transactions code, and thus these may interfere un-
predictably. However, these attempts at semantics for weak isolation are quite
complicated, and there is as yet no consensus on them.

By contrast, we assume a form of strong isolation, as described earlier. We
aim for a clean model that can highlight the issues which are specific to trans-
actions, and we specifically do not want to obfuscate them with difficult issues
that arise by introducing weak memory models, weak consistency, out-of-order
execution, and weak isolation. Such notions are somewhat orthogonal, and are
problematic semantically even in settings without transactions. Our goal is to
define an abstract, idealized, yet relevant, model of transactional programming
that could in principle serve as a foundation for verification. There are various
design choices in the implementation of a transactional memory framework (see
[15] for a taxonomy of choices), and our TSM model reflects several such choices.
For instance, our definition of nested transactions is a form of closed nested trans-
actions. We do not consider so-called open nested transactions, where an inner
transaction may commit while an outer one aborts (because we can not see any
sensible semantics for them, even in the single-process purely sequential setting).

Mechanisms other than transactions, such as locks, have been proposed to
enforce “atomicity” and have been studied from a verification point of view. For
instance, concurrent reactive programs where processes synchronize with locks
were studied in [19] where a custom procedure exploiting “atomicity” (based
on Lipton’s reduction) is used to simplify the computation of “summaries” for
such programs. Also, several verification problems are shown to be decidable
in [13] for a restricted class of programs where locks are nested. Several other
restrictions of concurrent pushdown processes for which verification problems
are decidable have also been identified (e.g., [9], among others). There are some
high-level similarity between these prior results and our results in Section 4, but
the details are substantially different due to the specifics of the TSM model.

Other related work discusses how to check the correctness of implementations
of transactional memory, based on lower level constructs, using testing ([16]) or
model checking ([10]). By contrast, we do not address the problem of analyzing
the correctness of implementations of transactional memory, but rather the correctness of transactional programs running on top of (correct) implementations.

Notions of serializability have been studied in database concurrency control for decades ([6]). However, there are subtle distinctions between the semantics of serializability in different setting. [4] systematically studied automata-based formalization of serializability and other related concepts. We formulate a clean and natural notion of stutter-serializability for TSMs, and show it is satisfied by them. The notion arose from our considerations of the TSM model, and does not appear to have been studied before in the literature.

2 Definition of Transactional State Machines

In this section we define Transactional State Machines (TSMs). The definition resembles that of (concurrent) boolean programs and (concurrent) extended recursive state machines (see, e.g., [5, 2, 3]), but with additional constructs for transactions. Our definition will use some standard notions (e.g., valuations of variables, expressions, types, etc.) which are defined formally in appendix A.1.

2.1 Syntax of TSMs

A Transactional State Machine $A$ is a tuple $A = \langle S, \sigma_{\text{init}}, (P_r)_{r=1}^n \rangle$, where $S$ is a set of shared variables, $\sigma_{\text{init}}$ is an initial valuation of $S$, and $P_1, \ldots, P_n$, are processes. Each process is given by $P_r = (L_r, \gamma_{\text{init}}, p_r, (A_r^i)_{i=1}^{k_r})$ where $L_r$ is a finite set of (non-shared) thread-local variables for process $r$, $\gamma_{\text{init}}$ is an initial valuation of $L_r$, $p_r \in \{1, \ldots, k_r\}$ specifies the index of the initial (main) procedure, $A_r^{p_r}$, for process $r$ (where runs of that process begin). The $A_r^i$’s are the procedures (or components in the RSM terminology) for process $r$. We assume that the first $d_r$ of these are ordinary and the remaining $k_r - d_r$ are transactional procedures. The two types of procedures have a very similar syntax, with the slight difference that transactional procedures have access to an additional abort node, $ab_i$. Specifically, each procedure $A_r^i$ is formally given by: $\langle N_r^i, en_r^i, ex_r^i, ab_r^i, \delta_r^i \rangle$, whose parts we now describe (for less cluttered notation, we omit the process superscript, $r$, when it is clear from the context):

- a finite set $N_i$ of nodes (which are control locations in the procedure)
- special nodes: $en_i, ex_i \in N_i$, known respectively as the entry node and exit node, and (only in the case of transactional components) also an abort node $ab_i \in N_i$.
- A set $\delta_i$ of edges, where an edge can be one of two forms:

\[3\] We note that these thread-local variables are used by all procedures running on the process. For simplicity, we do not include procedure-local variables, and we assume procedures take no parameter values and pass no return values. This is done only for clarity, and we lose nothing essential by making this simplification.
• **Internal edge:** A tuple \((u, v, g, \alpha)\). Here \(u\) and \(v\) are nodes in \(N_i\), \(g \in \text{BoolExp}(S \cup L_r)\) is a guard, given by a boolean expression over variables from \(S \cup L_r\) (see appendix A.1). \(\alpha \in \text{Assign}(S \cup L_r)\) is a (possibly simultaneous) assignment over these variables (again, see appendix A.1 for definition of assignments). We assume that \(u\) is neither \(\text{ex}_i\) nor \(\text{ab}_i\) (because there are no edges out of the exit or abort nodes), and that \(v\) is not the entry node \(\text{en}_i\). Intuitively, the above edge defines a possible transition that can be applied if the guard \(g\) is true, and if it is applied the simultaneous assignments are applied to all variables (all done atomically), and the local control node (i.e., program counter) changed from \(u\) to \(v\). The set of internal edges in procedure \(A_i\) is denoted by \(\delta^I_i\).

• **Call edge:** A tuple \((u, v, g, c)\). \(u\) and \(v\) are nodes in \(N_i\), \(g \in \text{BoolExp}(S \cup L_r)\) is a guard, \(c \in \{1, 2, \ldots, k\}\) is the index of the procedure being called. Again, we assume \(u \notin \{\text{ex}_i \cup \text{ab}_i\}\), and \(v \neq \text{en}_i\). Calls are either transactional or ordinary, depending on whether the component \(A_c\) that is called is transactional or not (i.e., whether \(c > d_r\) or \(c \leq d_r\)). Intuitively, a call edge defines a possible transition that can be taken when its guard \(g\) is true, and the transition involves calling procedure \(A_c\) (which of course involves appropriate call stack manipulation, as we’ll see). Upon returning (if at all) from the call to \(A_c\), control resumes at control node \(v\). The set of call edges in component \(i\) is denoted by \(\delta^C_i\).

### 2.2 Semantics of TSMs

A full formal semantics of TSMs is given in Appendix A.2 (due to space constraints). Here we give an informal description to facilitate intuition and describe salient features. TSMs model concurrent shared memory imperative procedural programs with bounded data and transactions. A configuration of an TSM consists of a call stack for each of the \(r\) processes, a current node (the program counter) for each process, as well as a universal valuation (or universal copy), \(U\), of shared variables. Crucially, during execution the “view” of shared variables may be different for different processes that are inside transactions. In particular, different processes, when executing inside transactions, will have their own local copies (valuations) of shared variables on their call stack, and will evaluate and manipulate those valuations in the middle of transactions, rather than the single universal copy.

A transaction keeps track (on the stack) of what shared variables have been used and written. If a shared variable is written by one of the processes inside the scope of one of the transactions, the universal copy \(U\) is not modified. Instead, an in-scope mutable copy of that shared variable is modified. The mutable copy resides on the stack frame of the innermost transaction on the call stack for that process. The first time a shared variable is read (i.e., used) inside a transaction, unless it was already written to in the transaction, its value is copied from \(U\) to a fixed local copy for that transaction and also to a separate mutable local copy, both on the stack. Thereafter, both reads and writes inside the transaction will be to this mutable copy.
At the end of a transaction, the written values will either be committed or aborted. The transaction is automatically aborted if a shared memory conflict has arisen, which is checked using an atomic compare-and-swap operation as follows. We compare the values of variables in the fixed copy of shared memory with the values of those same shared variables in the universal copy $U$, and if these are all equal, then for the written variables, we copy their valuations in the mutable copy on the stack frame to the universal copy $U$. If, on the other hand, the compare-and-swap fails, i.e., the compared values are not all equal then we abort the transaction, discard any updates to shared variables, pop the transactional stack frame and restore the calling context. (How this all works is described in detail in appendix A.2).

If we have nested transactions, and values are committed inside an inner nested transaction, then their effect will only be immediately visible in the next outer nested transaction (i.e., this follows the semantics of closed nested transactions), and the committed values will only be placed in the mutable copy of shared variables of the next outer transaction. Otherwise, if the inner transaction aborts, then its effect on shared variables is discarded before control returns to the calling context.

Again, see appendix A.2 for detailed semantics. We highlight here some other salient features of the semantics which will be pertinent in other discussions:

- The universal valuation $U$ is only updated upon a successful commit of outermost (non-nested) transactions, not of inner (nested) transactions.
- There are two distinct ways in which an abort can occur. One is an explicit abort, which occurs if a transaction reaches a designated abort node. The other is an automatic abort, carried out by the memory system due to conflicts with universal memory. (For nested transactions, the only possible abort is an explicit one, because no conflict is possible.)

3 Correctness: stutter-serializability

In this section we discuss a correctness property that TSMs possess. Informally, the correctness property relates to “atomicity” and serializability of transactions, but such notions have to be defined carefully with respect to the model. What we wish to establish is the following fact: if there exists any run $\pi$ of a TSM which witnesses a (possibly infinite) sequence of change to the universal copy (valuation) of memory, there must also exist a run $\pi'$ which witnesses exactly the same sequence of changes to the universal copy, but such that all transactions which start and which do not abort and do terminate in $\pi'$ must execute entirely serially without any interleaving of steps on other processes in the execution of the terminating transaction. Formally, this requires us to consider stutter-invariant temporal properties over atomic predicates that depend only on the universal valuation of shared variables in a state of the TSM, and stutter-equivalence (see appendix section A.3 for definitions).

We say a run $\rho$ of a TSM contains only serialized successful transactions if every transaction on any process in the run $\rho$ that starts and successfully
commits, executes serially without any interleaving of steps by other processes. In other words, the entire execution of each successful transaction occupies some contiguous sequence $\rho_i, \rho_{i+1} \ldots \rho_{i+m}$ in the run. For a run $\rho$ of $A$, let $\rho[U]$ denote a new word, over the alphabet of shared variable valuations, such that $\rho[U]$ is obtained from $\rho$ by retaining only the universal valuation of shared variables at every position of the run (i.e., replacing each state $\rho$ by the universal valuation in that state). We say that a TSM, $A$, is stutter-serializable if: for every run $\rho$ of $A$, there exists a (possibly different) run $\rho'$ of $A$ such that $\rho[U]$ is stutter-equivalent to $\rho'[U]$, and such that $\rho'$ contains only serialized successful transactions.

**Theorem 1.** All TSMs are stutter-serializable.

**Proof.** (Sketch) A proof is in the appendix. Here we sketch the basic intuition. If at the end of a non-nested transaction which is about to attempt to commit, the atomic compare-and-swap operation succeeds, then at exactly the point in “time” when the compare-and-swap operation executed, the values in the universal copy of shared variables used inside the transaction are exactly the same as the values that were read from the universal copy the first time these variables were encountered in the transaction. Each shared variable is read from the universal copy at most once inside any transaction. All subsequent accesses to shared variables are to the local mutable copy on the transactional stack frame. Consequently, since the values of shared variables are the only input to the transaction from its “environment” (i.e., from other processes), the entire execution of that transaction can be “delayed” and “rescheduled” in the same “infinitesimal time slot” just before the atomic compare-and-swap operation occurred, and the resulting effect of the transaction on the universal copy of memory after it commits would be identical (because it would have identical input). The only visible effect on the universal copy of memory during the run that this rescheduling has is that of adding or removing “stuttering” steps, because the rescheduled steps do not change values in the universal copy of shared memory.

Note that TSMs can reach new states due to transactions being aborted by the run-time environment due to memory conflicts. In other words, even aborted transactions have side effects. For instance, a TSM can use a thread-local variable to test/detect that its last (possibly nested) transaction was aborted, and take appropriate measures accordingly, including reaching new states that are reachable only following such an abort. This fact does not contradict the above correctness assertion about TSMs, because the correctness assertion does not rule out the possibility that in order for a certain feasible sequences of changes to universal memory $U$ to be realized some transactions might necessarily have to abort during the run. In general, it does not seems possible to devise a reasonable model of imperative-style transactional programs where transactions that are aborted will have no side effects. Anyway, there are good reasons not to want this. One useful consequence of side effects is that one can easily implement a “retry” mechanism in TSMs which repeatedly tries to execute the transaction until it succeeds. Various transactional memory implementations offer “retry” as a separate construct (see [15]).
4 Model Checking

It can be easily observed (via arguments similar to, e.g., [20]) that model checking for general TSMs, even with 2 processes, is at least as hard as checking whether the intersection of two context-free languages is empty. We thus have:

**Proposition 1.** Model checking arbitrary TSMs, even those with 2 processes, even against stutter-invariant LTL properties of shared memory is undecidable.

On the other hand, we show next that there is an interesting class of TSMs for which model checking remains decidable. Let the class of top-transactional TSMs be those TSMs with the property that the initial (main) procedure for every process makes only transactional calls (but inside transactions we can execute arbitrary recursive procedures). Let us call a TSM almost-top-transactional if one process is entirely unrestricted, but all other processes must have main procedures which make only transactional calls, just as in the prior definition.

**Theorem 2.** The model checking problem for almost-top-transactional TSMs against all stutter-invariant linear-time (LTL or $\omega$-regular) properties of (universal) shared memory is decidable.

**Proof.** Given a TSM, $A$, our first task will be to compute the following information. For each process $r$ (other than the one possible process, $r'$, which does not have the property that all calls in its main procedure are transactional) we will compute, for every transactional procedure, $A_c$ on process $r$, certain generalized summary paths. A generalized summary path (GSP) for a transactional procedure $A_c$ is a tuple $G = (\gamma_{\text{start}}, R, \gamma_{\text{finish}}, \text{status}, \sigma)$. $\gamma_{\text{start}}$ and $\gamma_{\text{finish}}$ are valuations of the thread-local variables $L_r$. status is a flag that can have either the value commit or abort. $\sigma$ is a partial valuation of shared variables, meaning it is a set of pair $(x,w)$ where $x$ is a shared variable and $w$ is a value in $x$'s domain (and there is at most one such pair in $\sigma$ for every shared variable $x$). $R = R_1, \ldots, R_d$ is a sequence of distinct partial valuations of shared variables, where furthermore, different $R_i$'s do not evaluate the same variable. In other words, for each shared variable $x$, there is at most one pair of the form $(x, w)$ in the entire sequence $R$. Such a sequence $R$ yields a partial valuation $\sigma_R = \bigcup_{i=1}^{d} R_i$ (and we shall need to refer both to the sequence $R$ and to $\sigma_R$).

We now define what it means for a GSP, $G$, to be valid for the transactional procedure $A_c$. Informally, this means that $G$ summarizes one possible terminating behavior of the transaction $A_c$ if it is run in sequential isolation (with no other process running). More formally, we call a GSP, $G = (\gamma_{\text{start}}, R, \gamma_{\text{finish}}, \text{status}, \sigma)$, valid for the transactional procedure $A_c$, if it satisfies the following property. Suppose a call to $A_c$ is executed in sequential isolation (i.e., with no other process running). Suppose, furthermore that in the starting state $\psi_0$ in which this call is made $\gamma_{\text{start}}$ is the valuation of thread-local variables $L_r$ on process $r$, and that the universal copy of shared memory $U$ is consistent with the partial valuation $\sigma_R$ (in other words it agrees with $\sigma_R$ on all variables evaluated in $\sigma_R$).
Then there exists some sequential run of $A_c$ from such a start state $\psi_0$ where during this run:

1. The sequence of reads of the universal copy of shared memory variables executed during the run corresponds precisely to the $d$ partial valuations $R_1, \ldots, R_d$. For example, if $R_3 = \{(x_1, w_1), (x_2, w_2)\}$, then the third time during the run in which the universal copy of shared memory is accessed (i.e., third time when shared variables are used that have not been used or written before) requires a simultaneous read\(^4\) of shared variables $x_1$ and $x_2$ from the universal copy $U$, and clearly the values read will be $w_1$ and $w_2$, because $U$ is by definition consistent with $\sigma_R$. (Note that $U$ does not change in the middle of the sequential execution of $A_c$, because it is run in sequential isolation, with no other process running.)

2. After these sequences of reads, the run of $A_c$ terminates in a state where the valuation of local variables is $\gamma_{\text{finish}}$ and either commits or aborts, consistent with the value of $\text{status}$.

3. Moreover, if it does commit, then the partial valuation of shared variables that it writes to the universal copy $U$ (via compare-and-swap) at the commit point is $\sigma$. (And otherwise, $\sigma$ is by default the empty valuation.)

Let $G_c$ denote the set of all valid GSPs for transactional procedure $A_c$. It is clear that for any transactional procedure, every GSP $G$ is a finite piece of data, and furthermore that there are only finitely many GSPs. This is because the universal valuation of every shared variable can be read at most once during the life of the transaction, and of course there are only finitely many variables, and each variable can have only finitely many distinct values.

**Lemma 1.** The set $G_c$ is computable for every transactional procedure $A_c$.

See appendix A.5 for a proof of the Lemma. We shall compute the set $G_c$ for every transactional procedure $A_c$ and use this information to construct a finite-state summary state-machine $B_r$, for every process $r$, which summarized that process’s behavior. We will also describe the behavior of the single unrestricted process $r'$ using a Recursive State Machine (RSM), $B'_{r'}$. We shall then use these $B_r$’s and $B'_{r'}$ to construct a new RSM $B = (\otimes_{r \neq r'} B_r) \otimes B'_{r'}$, which is an appropriate asynchronous product of all the $B_r$’s and $B'_{r'}$. The RSM $B$ essentially summarizes (up to stutter-equivalence) the behavior of the entire TSM with respect to shared memory. The construction of the $B_r$’s, $B'_{r'}$, and $B$ is described in appendix A.6.

It follows from the construction that $B$ has the following properties. For every run $\rho$ of the entire TSM, $A$, there is a a run $\pi$ of $B$ such that $\pi$ is stutter-equivalent to the restriction $\rho[U]$ of the run $\rho$ to its sequence of universal shared memory valuations. And likewise, for every run $\pi$ of $B$, there is a run $\rho$ of $A$ such that $\rho[U]$ is stutter-equivalent to $\pi$. (Again, see A.3 for definitions pertaining to stutter-equivalence.) Thus, once the RSM $B$ is constructed, we can use the model

\(^4\) Again, recall that the reason there may be simultaneous reads from universal shared variables is an artifact of the strong isolation assumption combined with our formulation of (potentially simultaneous) assignment statements.
checking algorithm for RSMs ([2]) on B to check any given stutter-invariant LTL, or stutter-invariant \( \omega \)-regular, property of universal shared memory of A.

We remark that the complexity of this model checking algorithm can be shown to be singly-exponential in the encoding size of the TSM, under a natural encoding of TSMs. (Note that TSMs are compactly encoded: they are a form of extended concurrent recursive state machines, with variables that range over bounded domains.)

Acknowledgements: Thanks to Jim Larus for several very helpful discussions and encouragement. Thanks to Tom Ball, Madan Musuvathi, Shaz Qadeer, and Serdar Tasiran for helpful discussions.

References

A Appendix

A.1 Basic definitions used in the TSM model

We require some preliminary definitions. Let $V$ denote a set of variables. Each variable, $v \in V$ has a type, $T(v) \in \text{Types}$. Each type, $t \in \text{Types}$, is associated with a domain, $D(t)$ (which may in general be infinite, but for analysis purposes we will assume it to be finite for all variables). In particular, we allow a boolean type, $\text{Bool}$, with domain $D(\text{Bool}) = \{T,F\}$. Let $\text{Expr}$ denote a set of expressions, which we assume are built up from variables, using operators and combinators. In particular, all variables are expressions. Each expression $\text{expr} \in \text{Expr}$ also has a type $T(\text{expr})$. For example, a boolean expression, $\text{bexpr}$, is any expression which has type $T(\text{bexpr}) = \text{Bool}$. We let $\text{BoolExp}(V)$ denotes the set of boolean expressions over variables $V$. The expressions of various types are typically given by grammars which describe how well-formed expressions are built up from variables of various types and using appropriate operators and constants. But we leave these details unspecified. We shall call the variables that appear in an expression its free variables. For a set of expressions $\text{Expr}'$, and a set of variables $V'$, we let $\text{Expr}'(V') \subseteq \text{Expr}'$ denote those expressions whose free variables are a subset of $V'$. For example, for an integer variable $x$, the following is a possible boolean expression: $(x \geq 5) \lor (x < -3)$. It has one free integer variable, $x$.

A valuation (or interpretation) of $V$ is a map, $\sigma : v \in V \mapsto d \in D(T(v))$. In other words, $\sigma$ maps every variable $v \in V$ of type $t$ to an element of the domain $D(t)$. It is assumed that every valuation $\sigma$ of the free variables $V'$ in a set of expressions $\text{Expr}(V')$ uniquely extends to a valuation for the expressions, i.e., to a map $\sigma : \text{expr} \in \text{Expr}(V') \mapsto d \in D(T(\text{expr}))$.

An assignment over $V$ has the form $[x_1, x_2, \ldots, x_k] := [\text{exp}_1, \text{exp}_2, \ldots, \text{exp}_k]$, where $x_j \in V$ are distinct variables, and for all $j$, $\text{exp}_j \in \text{Expr}(T(x_j))$. The semantics of assignments are defined over pairs $(\sigma_1, \sigma_2)$ of interpretations of $V$. Given an assignment $\alpha$ of the above form, we use the notation $\sigma_2 := \alpha(\sigma_1)$ to denote that (1) $\sigma_2(x_j) = \sigma_1(\text{exp}_j)$ for all $x_j$, and (2) $\sigma_1(y) = \sigma_2(y)$ for all other variables $y \in V \setminus \{x_1, x_2, \ldots, x_k\}$. In other words, apply assignment $\alpha$ to valuation $\sigma_1$ yields the new valuation $\sigma_2$. 
A.2 Semantics of TSMs

A configuration (i.e., state), $\psi$, of a TSM, $A$, is a tuple $\psi = \langle U, (stack_j, u_j, \gamma_j)_{j=1}^n \rangle$, where $U$, called the universal copy of shared memory, is a valuation of the set $S$ of shared variables. For each $j$, stack$_j$ is a stack, i.e., a (possibly empty) list of "stack frames". The contents of stack frames is described in detail later. $u_j$ is a node in some procedure of process $j$, which denotes the current control location (i.e., the program counter). We shall call this the current node of process $j$ in configuration $\psi$. $\gamma_j$ is a valuation of the thread-local variables $L_j$ on process $j$. The initial configuration of $A$ is $\psi_0 = \langle \sigma_{init}, (\emptyset, en_{pr}, \gamma_{pr})_{n=1}^n \rangle$. This starts each process at the entry point, $en_{pr}$ of its initial ("main") procedure, $A_{pr}$, with an empty call stack, and initializes the (universal) values of shared variables, and initializes the thread-local variables of each process. We shall now describe the global transition relation of a TSM by describing the effect of each kind of possible transition on the current configuration, including the stacks. When the TSM is in a given configuration (i.e., at a given state), a number of possible transitions can be enabled. In particular, a transition can be enabled because the guard, $g$, of some (internal or call) edge, $e$, of the TSM holds true in that configuration. We can categorize transitions basically into various different types: Calls, Internal transitions, (successful) Returns (which are Commits, in case the procedure is transactional), and (explicit or automatic) Aborts. We will consider each possible type of transition separately.

In order to simplify our notation when describing these, we have to make a natural adjustment to the interpretation of guards $g$ and assignments $\alpha$. Namely, guards (and assignments) define functions of valuations of variables, but which copies of those variables? Recall that there are potentially multiple copies of the same shared variable, $x$, including the universal copy as well as fixed/mutable copies on different stack frames. Different copies will be in-scope at different times during the execution. A guard $g(\bar{x})$ is only a function of the in-scope valuation of relevant variables, $\bar{x}$, and likewise an action $\alpha$ evaluates expressions over the in-scope copies of variables and assigns values to in-scope copies of variables. What copies of variables are in scope depends on the stack, as will become clear from the descriptions we shall give of different transitions.

Therefore, for simplifying notation, when the guard, $g$, of a (internal or call) edge on process $j$ is evaluated, we write $g$ as a boolean-valued function $g(U, stack_j, \gamma_j)$ of both the stack at process $j$ and the current universal valuation, $U$ of shared variables, as well as the current valuation $\gamma_j$ of thread-local variables. Likewise, we abuse notation somewhat and write the effect of an assignment $\alpha$ as $(U', stack_j', \gamma_j') := \alpha(U, stack_j, \gamma_j)$. Here the effect is again a mapping from the in-scope valuation of variables to a new in-scope valuation. The primed tuple $(U', stack_j', \gamma_j')$ on the left hand side reflects the effect that this mapping has on $(U, stack_j, \gamma_j)$.

We now consider, case by case, every possible type of transition. In every configuration, the stack, stack$_j$, of each process $j$, is a (possibly empty) sequence $SF_1, SF_2, \ldots, SF_m$ of stack frames. The stack grows from left to right, so $SF_m$ is the top-most stack frame.
– [Call]: If the current configuration is $\psi = \langle U, (stack_j, u_j, \gamma_j)_{j=1}^n \rangle$, and for some process $j$, which is currently at node $u_j$, there exists a call edge $e = (u_j, v, g, c)$, and the guard $g$ holds true in the current configuration $\psi$, i.e., $g(U, stack_j, \gamma_j) = \text{True}$, then an associated call transition is enabled in the current configuration. Assuming that $stack_j = SF_1, SF_2, \ldots, SF_m$, this transition will result in a new configuration $\psi'$, in which a new stack frame $SF_{m+1}$ will be pushed on to $stack_j$. Furthermore, in $\psi'$ the new node on process $j$ shall be updated to the entry node $en^j_c$ of component $A^j_c$ (i.e., the entry of the procedure that is called). The new stack frame, $SF_{m+1}$, will be as follows. If the procedure $A^j_c$ called by $e$ is transactional, then $SF_{m+1} := (e_{m+1}, TP_{m+1}, \text{Fixed}_{m+1}, \text{Mutable}_{m+1}, \text{Used}_{m+1}, Written_{m+1})$, otherwise, if it is non-transactional (ordinary), $SF_{m+1} := (e_{m+1}, TP_{m+1})$.

The contents of these tuples are as follows:

- $e_{m+1} = e = (u_j, v, g, c)$, is the call edge itself, which is stored on the stack frame for being able to restore the calling context upon returning/aborting.
- $TP_{m+1}$ is a transaction pointer, which points to the most recent transactional stack frame, if one exists, and is otherwise NULL. Specifically, if $e$ is a transactional call, i.e., if component $A^j_c$ is a transactional component, then $TP_{m+1} := m + 1$ (in other words, the transaction pointer points to its own stack frame as the latest transactional frame). If $e$ is not a transactional call, then: if the the stack $stack_j$ is not empty, then $TP_{m+1} := TP_m$, and if the call stack is empty then $TP_{m+1} := NULL$.

Furthermore, if the call $e$ is a transactional call, i.e., if $TP_{m+1} = m + 1$, then the stack frame $SF_{m+1}$ additionally contains, the following things:

- a set $\text{Fixed}_{m+1}$ of pairs of the form $(x, w)$, where $x$ is shared variable name and $w$ is a value in the domain of $x$, $D(T(x))$. The set $\text{Fixed}_{m+1}$ will hold a copy of those variables $x$ of universal shared memory whose values are used during the life of the outer-most transaction on the stack. This is the transaction which has an associated stack frame $SF_d$ which is transactional and such that there does not exist a smaller $d' < d$
such that $SF_d$ is transactional. The valuation in $Fixed_{m+1}$ will hold the values of these shared variables at precisely the time when they were first used during this outer-most transaction.

When the stack frame $SF_{m+1}$ is created, we initialize $Fixed_{m+1}$ as follows: if $TP_m \neq NULL$, then $Fixed_{m+1} := Fixed_{TP_m}$. Otherwise, (i.e., if $TP_m = NULL$, i.e. stack$_j$ is empty and $m = 0$) then $Fixed_{m+1} := \emptyset$.

- a set $Mutable_{m+1}$, which contains pairs $(x, w)$, where $x$ is a shared variable, and $w \in D(T(x))$ is a valuation of $x$. This mutable set will contain such pairs for the set of shared variables that have been used/written up to now during the life of the outer-most active transaction. Again, when the stack frame $SF_{m+1}$ is created, we initialize $Mutable_{m+1}$ just as with $Fixed_{m+1}$: if $TP_m \neq NULL$, then $Mutable_{m+1} := Mutable_{TP_m}$. Otherwise, (i.e., if $TP_m = NULL$ or stack$_j$ is empty and thus $m = 0$) then $Mutable_{m+1} := \emptyset$.
- a set $Used_{m+1}$ which containing the names of shared variables that have been used so far, again during the life of the outer-most active transaction. $Used_{m+1}$ is initialized in the same way as $Fixed_{m+1}$ and $Mutable_{m+1}$.
- a set $Written_{m+1}$ containing the names of shared variables that have been written so far, again, during the life of the outer-most active transaction. Again, $Written_{m+1}$ is initialized in the same way.

- **Internal Transition**: Internal transitions can occur when, in the current configuration $\psi = (U, (stack_j, u_j)_{j=1}^n)$, some process $j$ is at a control node $u_j$, and there exists some internal edge $e = (u_j, v, g, \alpha)$, such that the guard $g$ holds true in configuration $\psi$, i.e., $g(U, stack_j, \gamma_j) = True$. The transition, after evaluating the guard (and seeing that it holds true), applies the assignment $\alpha$, $(U', stack_j', \gamma_j') := \alpha(U, stack_j, \gamma_j)$. This of course, is done by first evaluating the relevant expressions on the in-scope copy of variables in $(U, stack_j, \gamma_j)$, and then assigning the values simultaneously to obtain the updated left hand side $(U', stack_j', \gamma_j')$. We then move to a new configuration where the current node on process $j$ is set to $v$.

Crucial assumption: it is crucially assumed that the combined effect of both first evaluating the guard $g$ and then carrying out the assignment $\alpha$ occurs atomically, meaning in particular that for an assignment of the form $[x_1, \ldots, x_d] := [expr_1, \ldots, expr_d]$, all assignments are made simultaneously and atomically. This is basically the **strong isolation** assumption.\(^6\)

When a variable, $x$ is read or updated by a guard or an assignment, if the variable is thread-local, then we read/update its value in the thread-local valuation $\gamma_j$. Then the variable $x$ is a shared variable, there are a number

---

\(^6\) It may appear too strong to assume that the sequence of *both* guard evaluation, and assignment, occur atomically. This is not, strictly speaking, necessary. We can easily rephrase the TSM model definition so that internal edges either have trivial guards which are always true, or trivial assignments which do nothing, without changing the expressive power of the TSM model. (For this, note that multiple edges, even on a single process, may be enabled at any given configuration. In other words, we have non-determinism available in the TSM model even at the level of a single process.)
of cases to consider, in order to decide what is the in-scope copy which is read/updated.

- if the transaction pointer of the top frame $TP_m = NULL$, i.e., we are not inside a transaction on process $j$, then we simply read, or write to, the actual universal copy of the shared variable $x$. In other words, we modify the universal valuation $U$.

- Otherwise, in the case when $TP_m = k \neq NULL$, i.e., we are inside a transaction on process $j$.

In this case, if the shared variable $x$ is read/used by the guard or assignment (i.e., occurs in one of the the right-hand-side expressions) then:

* if $x$ is not already on the $\text{Written}_k$ list of the current transaction frame (i.e., the stack frame $SF_k$ pointed to by $TP_m = k$), read its current value $w$ from the universal copy, $U$, and add $(x, w)$ to the set $\text{Mutable}_k$ of the current transaction frame. Furthermore, we also add $(x, w)$ to the $\text{Fixed}_k$ set of the current transaction frame. Finally, we add the variable $x$ to $\text{Used}_k$.

* if $x$ is already in the set $\text{Written}_k$, then simply look up and use its current value in $\text{Mutable}_k$. (A single valuation $(x, w)$ will already be in $\text{Mutable}_k$, because $x$ was already written to during the life of this transaction frame.)

If a variable $x$ is updated with a value $w$, i.e., $x$ is an assignment assigns $x$ some expression which evaluates to $w$, then:

* if $x$ isn’t already in $\text{Written}_k$, add it to $\text{Written}_k$.

* update the value of $x$ in $\text{Mutable}_k$. In other words, put the pair $(x, w)$ to $\text{Mutable}_k$, overwriting any earlier pair $(x, w’)$ in $\text{Mutable}_k$ if such a pair exists.

- [Return]: Such a transition is enabled if in the current configuration, on some process $j$, the current node is an exit node, i.e., $u_j = \text{ex}_j$. There are two cases to consider:

  - Return from an ordinary (non-transactional) procedure: In this case the topmost stack frame, $SF_m$ on stack $j$ is not a transactional frame, i.e., $TP_m \neq m$. In this case, for effecting this transition, we pop the stack frame $SF_m$, and we update the current node on process $j$ to the target $v$ of the call edge $e_m = (u, v, g, c)$ which is also available in $SF_m$.

  - Return (either committed or automatically aborted) from a transactional procedure: In this case, $TP_m = m$. If the current transaction frame is not the only transaction frame remaining (meaning $TP_{m-1} = d \neq NULL$), then we treat this exit similarly to a non-transactional procedure exit, except that we move the values of the fixed and mutable valuations of shared variables to their fixed and mutable copy in the next outer transaction frame (overwriting older values if necessary), and we do likewise with the used/written sets. In other words, given that $TP_{m-1} = d$, we assign $\text{Fixed}_d := \text{Fixed}_m$, $\text{Mutable}_d := \text{Mutable}_m$, $\text{Used}_d := \text{Used}_m$, and $\text{Written}_d := \text{Written}_m$.

Note that what this means is that there is never an automatic run-time abort of a nested transaction. There are only explicit aborts of nested transactions. Such aborts
If, on the other hand, this transaction frame is the only remaining one on the call stack (i.e., $TP_{m-1} = NULL$), we perform the following 1-step compare-and-swap: Pop the stack frame $SF_m$, and use the call edge $e_m$ which is stored on $SF_m$ to restore the context and assign a new node for process $j$.

- If, on the other hand, the compare fails, i.e., values in the universal copy and fixed copy on the stack of shared memory are not the equal then do exactly the same as done by an explicit abort operation, which is specified next.

  – **[Abort]:** An abort transition can occur, either because an explicit abort node is encountered and terminates a transaction, meaning in the current configuration some process $j$ is at an abort node, i.e., $u_j = ab_j$, or it can also occur because (as described at the end of the previous [Return] case), the compare-and-swap at the end of a transaction failed, resulting in a memory conflict, in which case the run-time system will do an automatic abort.

  In both cases, we do the following. If the abort is being carried out on process $j$, and the top stack frame on $stack_j$ is $SF_m$ (this is necessarily a transactional frame) then we pop $SF_m$ off $stack_j$, and we restore the new node (just as in the “Return” case).

  We also do the following important thing: if this was a nested transaction, meaning that $TP_{m-1} = d \neq NULL$, then we must update $Fixed_d$, $Mutable_d$, and $Used_d$ with the new values of shared variables that were read for the first time (in the context of transaction frame $SF_d$) inside this nested transaction $SF_m$. We do this update as follows: For every pair $(x, w) \in Fixed_m$ such that there does not exist any valuation $(x, w')$ of variable $x$ in $Fixed_d$, add $(x, w)$ to $Fixed_d$, and likewise add $(x, w)$ to $Mutable_d$ and add $x$ to $Used_d$.

  This updating is done so that these shared variables, whose universal copy has already been read once inside the context of the transaction frame $SF_d$ (in fact inside $SF_m$) will not be read again.

This completes the description of the semantics of TSMs. As usual, the set of configurations, $Q$, and the transition relation, $\Delta$, of a TSM, $A$, together define an ordinary (and in general infinite-state) transition system $T_A = (Q, \Delta)$. A run of $A$ is a (finite or infinite) sequence $\rho = \rho_0 \xmapsto{e_0} \rho_1 \xmapsto{e_1} \rho_2 \ldots$, where $\psi_0$ is the initial configuration and for all $i$, $\psi_i \in Q$ and $(\psi_i, e_i, \psi_{i+1}) \in \Delta$, where the event $e_i$ is a specific edge, return, commit, or abort, operation that is enabled in state $\rho_i$ and such that executing it yields state $\rho_{i+1}$. For a run $\rho$, we will use $\rho(i, j)$, $i \leq j$, to denote the segment $\rho_i \ldots \rho_j$ of the run. We refer to segments $\rho(0, j)$ as initial segments of $\rho$.

\[\text{can simply be viewed as exception handling mechanisms, which allow the effect of the nested transaction on shared memory to be “erased”. However, it should be noted that, inevitably, the nested transaction can nevertheless have “side effects” even if aborted. In particular, the nested transaction may have updated thread-local variables. Such side effects can easily be used to program, e.g., an explicit retry mechanism for transactions in this model, so separate constructs for retry are not needed.}\]
A.3 Stutter-invariance

Stutter-invariance was first considered in [14]. For a (finite or infinite) words $w = w_0w_1w_2\ldots$ over an alphabet $\Sigma$, with $w_i \in \Sigma$ for all $i$, and a mapping $f : \mathbb{N} \rightarrow \mathbb{N}^+$, from natural numbers to positive natural numbers, let $w^f = w_0^f w_1^f w_2^f \ldots$, where, as usual, $w_i^n$ denotes the concatenation of $n$ copies of the letter $w_i$. Recall that a language $L \subseteq \Sigma^\infty = \Sigma^\omega \cup \Sigma^*$ over an alphabet $\Sigma$ is stutter-invariant iff for all (finite or infinite) words $w = w_0w_1w_2\ldots$, and for all mappings, $f : \mathbb{N} \rightarrow \mathbb{N}^+$, $w \in L$ iff $w^f \in L$. Two words $w$ and $w'$ are stutter-equivalent if they agree on all stutter-invariant properties, i.e., for any stutter-invariant language $L$ they are either both contained in $L$ or both not. LTL formulas without the next operator express precisely all stutter-invariant LTL languages, and there are similarly easy syntactic restrictions of B"uchi automata for describing all stutter-invariant $\omega$-regular properties (see [18, 11]).

A.4 Proof of Theorem 1

**Theorem 1.** All TSMs are stutter-serializable.

**Proof.** Given any run $\rho$ of a TSM, $A$, let $\rho(0, j) = \rho_0 \ldots \rho_j$ be an initial segment of $\rho$ such that $\rho_j$ is the state just after a transaction on some process $k$ terminates successfully and commits. Let the transitions $\rho_i \xrightarrow{e_i} \rho_{i+1}$, $\rho_i \xrightarrow{e_i}$ $\rho_{i+1} = \rho_j$ be precisely those transitions of the run associated with the execution of a particular transaction on process $k$ from start to finish (i.e., from just before the transaction call to just after its successful commit).

Note that the only transition among these that can possibly change the universal valuation is the final (commit) transition $\rho_{m-1} \xrightarrow{e_m} \rho_m$.

We now inductively construct a new run $\rho'$, by induction on the length $j$ of such initial segments of $\rho$ whose last state is the state just after a successful commit of a transaction. Given such an initial segment, we delay the events $e_1, \ldots, e_{m-1}$ and reschedule them just before $e_m$. In other words, the new event sequence will look as follows:

$e_1e_2 \ldots e_i 1e_{i} e_{i+1} + e_{i+2} \ldots e_{i} e_{i+1} e_{i+2} + 1 \ldots e_{i=1} e_{i=2} + 1 \ldots e_{i=m-1} e_{i=1} e_{i=2} \ldots e_{i=m}$.

We claim that these events remain enabled after this reordering and generate a valid initial segment of a run, $\rho'(0, j) = \rho'_0 \ldots \rho'_j$, and that furthermore $\rho'(0, j) \cup B$ is stutter-equivalent to $\rho(0, j) \cup B$, and finally that the last state $\rho'_j$ is exactly equal to $\rho_j$. The reasons these claims hold was basically described in the sketch proof in the body of the paper: since the only communication between processes is through the universal copy of shared memory, no event on processes other than $k$ is disabled because of the delay of the non-committing events inside this transaction on process $k$, and furthermore, all the events in the transaction will remain enabled after the delay because, since the compare-and-swap of the commit operation $e_{m-1}$ succeeded in the original run, the universal values of all
shared variables remain the same as they were when each read inside the transaction took place. Note that stuttering of the universal values can be introduced (or removed) by this reordering of events, but the final states are identical, because every process’ view of the initial segment \( \rho'(0,j) \) is exactly the same as \( \rho(0,j) \). By induction on the length of initial segment of \( \rho \) which end in a successful commit, we can convert longer and longer initial segments, and thus the entire run \( \rho \), into a new run \( \rho' \) such that \( \rho'[\mathcal{U}] \) is stutter-equivalent to \( \rho[\mathcal{U}] \) and such that \( \rho' \) contains only serialized successful transactions. (Note that there is no danger that an event will be indefinitely delayed by this reordering, because the only events that are delayed in the reordering are non-commit events of a transaction which commits successfully, and they are only delayed a finite number of steps until just prior to the successful commit.) □

A.5 Proof of Lemma 1

**Lemma 1.** The set \( G_c \) is computable for every transactional procedure \( A_c \).

**Proof.** Recall that there are only finitely many possible GSPs. Moreover, even though a transaction may internally be a recursive procedure, since we are considering sequential executions of the transaction in isolation from all other processes we can treat the transaction just like a standard Recursive State Machine (RSM) (or, equivalently, Pushdown System) and use the well known algorithms for model checking them (see [2]) in order to determine whether an execution corresponding to each given generalized summary path is possible (i.e., valid) or not, as follows.

Specifically, for a given GSP, \( G \), we can use an LTL property to describe the possible sequence of (simultaneous) atomic reads corresponding to \( G \). Intuitively, we can use an LTL formula with nested untils to describe a sequence like the following: “the first shared variable that is read (atomically) is \( x_{i_1} \) with a value read of \( w_{i_1} \), and thereafter there are no reads of unused shared variables until \( x_{i_2} \) and \( x_{i_3} \) are read simultaneously, with values \( w_{i_2} \) and \( w_{i_3} \), and thereafter .....”. Furthermore, the LTL formula corresponding to \( G \) can also describe the values according to \( G \): of thread-local variables at the beginning and end of the execution, the values of shared variables that are committed (if any) to \( \mathcal{U} \) at the end, and whether the transaction did indeed commit or abort.

Model checking such an LTL formula for \( G \) against the Recursive State Machine corresponding to \( A_c \) will decide whether the given GPS, \( G \), is valid or not. We can do this check separately for every possible GPS, to compute the set \( G_c \) of valid ones. We can do so because there are only finitely many GSPs. (There are in fact more efficient ways to compute the valid GSPs in an aggregate way, using standard techniques from program analysis, but we don’t need these for the result.) □

A.6 Construction of \( B_r \), \( B'_r \), and \( B \), in the proof of Theorem 2.

We now describe how the finite summary state-machine, \( B_r \), for process \( r \), discussed in the proof of Theorem 2, can be constructed. A state of \( B_r \) contains a
valuation of local memory, a partial valuation of universal shared memory (the reason why it is partial will be made clear shortly), and a control location (control node) which corresponds either to a control node of the main procedure of process $r$, or to some auxiliary control nodes which are needed as intermediate states in the middle of mimicking the execution of generalized summary paths. Also, the states having such auxiliary control nodes will additionally need to keep an extra fixed partial valuation of shared memory, as described below.

Intuitively, the state machine $B_r$ mimicks the execution of the “main” procedure, $A_{pr}$ of process $r$, except recall that the main procedure can only make transactional calls. Suppose an enabled call $e = (u, v, g, c)$ to a transaction, $A_c$, is encountered, in a state where the thread-local valuation and (partial) universal valuations of variables are $\gamma$ and $\sigma'$. Rather than actually executing the call $e$ and maintaining a call stack, $B_r$ non-deterministically “guesses” some valid GSP, $G = (\gamma_{\text{start}}, R, \gamma_{\text{finish}}, \text{status}, \sigma) \in \mathcal{G}_c$ such that $\gamma_{\text{start}} = \gamma$. (If no such valid GSP exists, then that means no halting behavior of this transaction is possible if it is executed starting in the memory “environment” given by $\gamma$. In such a case, $B_r$ effectively “hangs” as a process.) If such a GSP, $G$ does exist, and if $B_r$ has guessed $G$ non-deterministically, it then goes through a sequence of $d$ intermediate auxiliary states, which reflect the sequence of $d$ (potentially simultaneous) reads $R = R_1, \ldots, R_d$ that occur in $G$, and the partial valuations of shared memory that they define. The intermediate auxiliary state associated with read $R_i$ has partial valuation $R_i$ of shared memory These intermediate states also retain another fixed partial valuation. Namely, the $j$'th intermediate state in the sequence contains fixed partial valuation $F_j = \bigcup_{i=1}^{j} R_i$. Finally, at the end of the auxiliary sequence we update local valuation $\gamma$ to $\gamma_{\text{finish}}$, and update the partial valuation of universal shared memory to a new valuation $\sigma''$, consistent with $\text{status}$ and $\sigma$, as follows. If $\text{status} = \text{commit}$, then we non-deterministically either set $\sigma'' := \sigma$ or $\sigma'' := \emptyset$, and if $\text{status} = \text{abort}$ then $\sigma'' = \emptyset$. The control node is then set to $v$, the return point of the call. Transitions of $B_r$ are labeled by either the edge $e$ of the TSM that enabled them, or by $\text{commit}$, $\text{automatic-abort}$, $\text{manual-abort}$ if it is a transaction’s commit step, automatic abort step, or manual abort step, respectively, or by $\text{return}$ if the transition corresponds to a return from an ordinary procedure call. If the transition corresponds to an intermediate auxiliary step within a transaction, it has a special label $\text{auxiliary}$. This completes our description of $B_r$. (We forgo more formal notation defining the states and transitions of $B_r$, which is rather tedious.) We can readily observe that since there is only finite data, and since

---

8 Such a call is enabled if we are at control node $u$ and the guard $g$ is satisfied by the thread-local valuation $\gamma$ together with any full shared valuation $\sigma^*$ which extends $\sigma'$, i.e., such that $\sigma' \subseteq \sigma^*$.

9 The two possibilities here correspond to either a successful commit or an automatic abort by the run-time system due to a memory conflict. Later in the construction of the asynchronous product $B$ this non-deterministic choice is resolved by checking consistency with the universal shared valuations of other processes.

10 In the final transition out of the auxiliary sequence we discard the final fixed partial valuation $F_d$. $F_d$ is used later in the construction of $B$. 

the length $d$ of such generalized summary paths is bounded, the number of states of $B_r$ is finite.

Construction of the single RSM, $B_r'$, corresponding to the unrestricted process $r'$ is easier, because we don’t have to incorporate summaries for it. We can retain all the stack processing done on one process of the TSM (including that done for transactions) on the call stack of the RSM, $B_r'$. We omit a detailed description of $B_r'$ here. (The only minor subtlety to observe is that states of $B_r'$ are also labeled by appropriate partial valuations of the universal copy of shared memory, as in the case of the summary state-machines $B_r$.)

Armed with the summary state machines $B_r$ for all “top-transactional” processes, together with the RSM, $B_r'$, corresponding to the single unrestricted process, we can construct the asynchronous product RSM $B = (\otimes_r B_r) \otimes B_r'$.

Informally, this asynchronous product mimicks the interleaved execution of all these processes when they run concurrently, including the possible interleaving of steps of other processes in the middle of auxiliary steps associated with a transaction running on one of the processes. A state $s$ of the asynchronous product $B$ consists of a tuple of process states, one for each process, such that the partial valuations of universal shared memory in all of the process states in the tuple are consistent with each other (meaning no two of them give different values to the same variable). A state $s$ of $B$ also contains a universal partial valuation $\sigma_s$ which is the union of the (consistent) universal partial valuations of each process. Transitions of $B$ are determined by transitions of individual processes: they transform the process state for that process only, together with the universal partial valuation $\sigma_s$. A transition on process $i$ is enabled at a state $s$, with process state $s_i$ for process $i$, under the following conditions. If that process’s transition is labeled by an edge $e$, then it is enabled if there is an extension of the thread-local (partial) valuation of that process’s state, $s_i$, and an extension of the combined universal (partial) valuation $\sigma_s$ of all processes, which satisfies the guard of edge $e$.11 Similarly, if that process’s transition is labeled by return or labeled auxiliary, then it is always enabled. A commit step is enabled if the fixed valuation $F_d$ in the process state $s_i$ is consistent with the universal valuation $\sigma_s$. By contrast, a automatic-abort step is enabled if the fixed valuation $F_d$ is inconsistent with $\sigma_s$. (This ensures that commit steps only occur if there is a successful compare-and-swap at the end of the transaction.) Finally, manual-abort steps are always enabled. This completes our description of $B$ (we forgo a more formal description.) It is not hard to see that since there is only one unrestricted process, $B$ can itself be viewed as an RSM, with the call stack of $B_r'$ now serving also as the call stack of the entire product RSM, $B$.

11 Of course, since we only keep states $s$ where partial valuations of universal memory are consistent, this also implicitly implies that the transition is only enabled if the valuations of shared memory that it results in, based on what variables are written to, is consistent between different processes.