Verifying a Quantitative Relaxation of Linearizability via Refinement

Kiran Adhikari

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Computer Engineering

May 6, 2013

Blacksburg, Virginia

Keywords: Quasi Linearizability, Refinement, Model Checking

Verifying a Quantitative Relaxation of Linearizability via Refinement

Kiran Adhikari

(ABSTRACT)

Concurrent data structures have found increasingly widespread use in both multicore and distributed computing environments, thereby escalating the priority for verifying their correctness. The thread safe behavior of these concurrent objects is often described using formal semantics known as *linearizability*, which requires that every operation in a concurrent object appears to take effect between its invocation and response. *Quasi linearizability* is a quantitative relaxation of *linearizability* to allow more implementation freedom for performance optimization. However, ensuring the quantitative aspects of this new correctness condition is an arduous task. We propose the first method for formally verifying quasi linearizability of the implementation model of a concurrent data structure. The method is based on checking the refinement relation between the implementation model and a specification model via explicit state model checking. It can directly handle multi-threaded programs where each thread can make infinitely many method calls, without requiring the user to manually annotate for the linearization points. We have implemented and evaluated our method in the PAT model checking toolkit. Our experiments show that the method is effective in verifying quasi linearizability and in detecting its violations.

Dedication

I dedicate this thesis to my grandfather, my parents and my brother.

Acknowledgment

I extend my gratitude to my advisor, Professor Chao Wang for his consistent guidance and continuous support throughout my research. His advice and encouragement have been extremely valuable in helping me complete my thesis. I would also like to thank Dr. Michael Hsiao and Dr. Patrick Schaumont for serving as members of my thesis committee.

I sincerely thank my seniors Sanjay Basnet, Gyanendra Shrestha, Nabin Jnawali, Ram Thapa and Bipin Jnawali for their consistent encouragement. For all my friends, including Diwas, Suju, Krishna, Keshab, Prakash and Suwarna, in particular thank you for your direct and indirect support and your company throughout the years. I am also deeply grateful to James Street, whose help was very beneficial to this thesis and I really enjoyed working together with him.

I am deeply indebted to my beloved grandfather, my parents and all the other family members, for their endless love and support. They have always been there for me, through thick and thin.

Kiran Adhikari

May 6, 2013

Contents

List of Figures

List of Tables

List of Abbreviations

BFS	Breadth-First Search
DFS	Depth-First Search
PAT	Process Analysis Toolkit
CSP	Communicating Sequential Processes
LTS	Labeled Transition System
QF	Quasi Factor

Chapter 1

Introduction

Concurrent programming has engulfed the broad horizon of both the multi-core and distributed computing environments. As Moore's law starts to pay little dividend in improving the performance of sequential applications, programmers need to adopt the concurrent programming paradigm to help reduce the memory contention and improve the scalability and throughput of the applications significantly. This has led to the increasingly widespread use of concurrent data structures in domains spanning from embedded computing to distributed systems. However, it is often difficult to build and test the thread safe components for the concurrent programming environment without placing undue synchronization overhead. The often large number of thread interleavings, along with the subtle interactions of concurrent operations, makes it difficult to obtain the correct implementations. This has escalated the priority for verifying the correctness of these concurrent data structures.

Over the past two decades, researchers have focused on using linearizability as a correctness condition for concurrent data structures. A concurrent data structure is linearizable if each of its operations or method calls appears to take effect instantaneously at some point in time between its invocation and response. Although being linearizable does not necessarily ensure the full-fledged correctness of the implementation, linearizability violations are clear indicators that the implementation is buggy. In this sense, linearizability serves as a useful criterion for implementing concurrent data structures. However, ensuring linearizability of highly concurrent data structures is a difficult task, due to the subtle interactions of concurrent operations and the often astronomically many interleavings.

Although linearizability is non-blocking[10], it often imposes unnecessarily tight synchronization requirement on the implementation and therefore limit the performance and scalability. Realizing that a more relaxed correctness condition than linearizability suffices in many applications, Afek *et al.* [1] have defined a new notion called *quasi linearizability*. Quasi linearizability is a quantitative relaxation of linearizability [9, 12, 17] to allow for more flexibility in how the data structures are implemented. While preserving the basic intuition of linearizability, quasi linearizability relaxes the semantics of the data structures to achieve increased runtime performance. For example, when implementing a queue for a task scheduler in a thread pool, it is often the case that we do not need the strict first-in-first-out (FIFO) semantics; instead, we may want to allow the dequeue operations to be overtaken occasionally if it helps improving the runtime performance. The only requirement here is that such out-of-order execution should be bounded by a fixed number of steps. Similarly, when implementing data caching in web applications, we may not need the strict semantics of standard data structures, since getting stale data occasionally is acceptable as long as the delay is bounded. Furthermore in distributed systems, the counter for generating unique identifiers may also be allowed to return out-of-order values occasionally.

1.1 Motivation

Despite the advantages of quasi linearizability and its rising popularity (e.g., [9, 12, 17]), such relaxed consistency property is difficult to test and verify. Inspit of being an important property of concurrent data structures, there does not yet exist any effective way of formally verifying this relaxed correctness condition.

Although there is a large body of work on formally verifying standard linearizability, for example, the methods based on model checking [15, 14, 24, 5], runtime verification [4], and mechanical proofs [22, 23], they cannot directly verify quasi linearizability. Because of the inherent non-determinism in the quasi linearizable data structures, these traditional verification methods do not have the capability of checking teh quasi linearizability semantics. In addition to the requirement of covering all possible interleavings of concurrent events, one needs to accurately analyze the quantitative aspects of these interleavings to verify quasi linearizability. Another reason why formal verification methods for quasi linearizability are lacking is because quasi linearizability is relatively new and still under development.

In this thesis, we fill the gap by developing an effecient method to verify this quantitative relaxation of the linearizability for concurrent data structures.

1.2 Contribution

We propose the first automated method for formally verifying *quasi* linearizability in the implementation models of concurrent data structures. There are several technical challenges. First, since the number of concurrent operations in each thread is unbounded, the execution trace of a multi-threaded program that uses the concurrent data structures may be infinitely long. This precludes the use of existing methods such as LineUp [4] because they are based on checking permutations of finite length execution histories. Second, since the method needs to be fully automated, we do not assume that the user will provide hints or annotate the linearization points of each method. This precludes the use of existing methods that are based on either user guidance (e.g., [22, 23]) or annotated linearization points (e.g., [24]).

To overcome these challenges, we rely on explicit state model checking. That is, given an implementation model M_{impl} and a specification model M_{spec}, we check whether the set of execution traces of M_{impl} is a subset of the execution traces of M_{spec}. Toward this end, we extend a classic refinement checking algorithm so that it can check for the newly defined *quantitative relaxation* of standard refinement relation.

Consider a quasi linearizable queue as an example. Starting from the pair of initial states of a FIFO queue specification model and its quasi linearizable implementation model, we check whether all subsequent *state transitions* of the implementation model can match some subsequent *state transitions* of the specification model. To make sure that the verification problem remains decidable, we bound the capacity of the data structure in the model, to ensure that the number of states of the program is finite.

We have implemented the new method in the Process Analysis Toolkit (PAT) [20], which is a model checker for analysing concurrent systems. PAT provides the basic infrastructure for parsing and analyzing the specification and implementation models written in a process algebra language that resembles Tony Hoare's Communicating Sequential Processes (CSP) [11]. Our new method is implemented as a module in PAT, and is compared against the existing module for checking standard refinement relation. Our experiments on a set of popular concurrent datastructures, such as queues, stacks or priority queues show that the new method is effective in detecting subtle violations of quasi linearizability. When the implementation model is indeed correct, our method can also generate the proof of correctness quickly.

To sum up, this thesis make the following contributions:

- Propose the first method for formally verifying quasi linearizability of concurrent data structures. This is accomplished by designing a new algorithm for checking a relaxed version of the refinement relation between the implementation and specification models.
- Implement the new method in a software tool called PAT and demonstrate its effectiveness on a set of quasi linearizable concurrent data structure examples including queues, stacks, and priority queues.

1.3 Organization

The remainder of this thesis is organized as follows.

Chapter 2 provides the overview of the notations and review the existing refinement checking algorithm for verifying the standard linearizability.

Chapter 3 introduces the overall flow of our new relaxed linearizability checking method. This chapter also presents a manual approach for verifying quasi linearizability based on the standard refinement checking algorithm. This approach is proved to be labor intensive and error prone, therefore motivating us to design a fully automated method.

Chapter 4 presents the fully automated method, which is based on a new algorithm for checking the relaxed refinement relation. This chapter also presents the experimental results.

Chapter 5 gives the conclusions and outlines the future works.

Chapter 2

Background

In this chapter, the notion of linearizability and quasi linearizability are defined by using the refinement relations between an implementation model and the specification model. We also present the process algera semantics of the models in relation to the CSP. Finally, we introduce the standard refinement checking algorithm as an approach to verify the linearizability. The method relies on model checking of finite state systems specified as concurrent processes with shared variables and take advantage of a new trace refinement checking to verify the relaxed version of linearizable data structures.

Given a two labeled transition systems, the following section explains the formal semantics defining standard and quasi linearizability, and review an existing algorithm for checking the refinement relation between those two labeled transition systems.

2.1 Linearizability

Linearizability [10] is a safety property of concurrent systems, over sequences of actions corresponding to the invocations and responses of the operations on shared objects. We begin by formally defining the shared memory model.

Definition 1 (System Models). *A shared memory model $\mathcal{M}$ is a 3-tuple structure $(O, init_O, P)$, where O is a finite set of shared objects, $init_O$ is the initial valuation of O, and P is a finite set of processes accessing the objects.* $\square$

Every shared object has a set of states. Each object supports a set of *operations*, which are pairs of invocations and matching responses. These operations are the only means of accessing the state of the object. A shared object is *deterministic* if, given the current state and an invocation of an operation, the next state of the object and the return value of the operation are unique. Otherwise, the shared object is *non-deterministic*. A *sequential specification*[1] of a deterministic (resp. non-deterministic) shared object is a function that maps every pair of invocation and object state to a pair (resp. a set of pairs) of response and a new object state.

An execution of the shared memory model $\mathcal{M} = (O, init_O, P)$ is modeled by a history, which is a sequence of operation invocations and response actions that can be performed on O by processes in P. The behavior of $\mathcal{M}$ is defined as the set, H, of all possible histories together. A history $\sigma \in H$ induces an irreflexive partial order $<_\sigma$ on operations such that $op_1 <_\sigma op_2$ if the response of operation op_1 occurs in σ before the invocation of operation op_2. Operations in σ that are not

[1]More rigorously, the sequential specification is for a *type* of shared objects. For simplicity, however, we refer to both actual shared objects and their types interchangeably in this paper.

related by $<_\sigma$ are concurrent. A history σ is $sequential$ iff $<_\sigma$ is a strict total order.

Let $\sigma|_i$ be the projection of σ on process p_i, which is the subsequence of σ consisting of all invocations and responses that are performed by p_i in P. Let $\sigma|_{o_i}$ be the projection of σ on object o_i in O, which is the subsequence of σ consisting of all invocations and responses of operations that are performed on object o_i. Every history σ of a shared memory model $\mathcal{M} = (O, init_O, P)$ must satisfy the following basic properties:

- **Correct interaction:** For each process $p_i \in P$, $\sigma|_i$ consists of alternating invocations and matching responses, starting with an invocation. This property prevents *pipelining*[2] operations.

- **Closedness[3]:** Every invocation has a matching response. This property prevents *pending* operations.

A sequential history σ is *legal* if it respects the sequential specifications of the objects. More specifically, for each object o_i, there exists a sequence of states s_0, s_1, $s_2, \ldots$ of object o_i, such that s_0 is the initial valuation of o_i, and for all $j = 1, 2, \ldots$ according to the sequential specification (the function), the j-th invocation in $\sigma|_{o_i}$ together with state s_{j-1} will generate the j-th response in $\sigma|_{o_i}$ and state s_j. For example, a sequence of read and write operations of an object is *legal* if each read returns the value of the preceding write if there is one, and otherwise it returns the initial

[2]Pipelining operations mean that after invoking an operation, a process invokes another (same or different) operation before the response of the first operation.

[3]This property is not required in the original definition of linearizability in [10]. However adding it will not affect the correctness of our result because by Theorem 2 in [10], for a pending invocation in a linearizable history, we can always extend the history to a complete one and preserve linearizability. We include this property to obviate the discussion for pending invocations.

value.

Given a history σ, a *sequential permutation* π of σ is a sequential history in which the set of operations as well as the initial states of the objects are the same as in σ.

Definition 2 (Linearizability). *Given a model $\mathcal{M} = (O = \{o_1, \ldots, o_k\}, init_O, P = \{p_1, \ldots, p_n\})$. Let H be the behavior of $\mathcal{M}$. $\mathcal{M}$ is linearizable if for any history σ in H, there exists a sequential permutation π of σ such that*

1. *for each object o_i ($1 \leq i \leq k$), $\pi|_{o_i}$ is a legal sequential history (*i.e.*, π respects the sequential specification of the objects), and*

2. *for every op_1 and op_2 in σ, if $op_1 <_\sigma op_2$, then $op_1 <_\pi op_2$ (*i.e.*, π respects the run-time ordering of operations).* $\square$

Linearizability can be equivalently defined as follows. In every history σ, if we assign increasing time values to all invocations and responses, then every operation can be shrunk to a single time point between its invocation time and response time such that the operation appears to be completed instantaneously at this time point [16, 3]. This time point is called its *linearization point*.

2.2 Quasi Linearizability

For two histories σ and σ' such that one is the permutation of the other, we define their distance as follows. Let $\sigma = e_1, e_2, e_3, \ldots, e_n$ and $\sigma' = e'_1, e'_2, e'_3, \ldots, e'_n$. Let $\sigma[e]$ and $\sigma'[e]$ be the indices of the event e in histories σ and σ', respectively. The distance between the two histories, denoted $\Delta(\sigma, \sigma')$, is defined as follows:

$$\Delta(\sigma, \sigma') = max_{e \in \sigma}\{|\sigma'[e] - \sigma[e]|\} \ .$$

In other words, the distance between σ and σ' is the maximum distance that an event in σ has to move to arrive at its position in σ'.

While measuring the distance between two histories, we often care about only a subset of method calls. For example, in a concurrent queue, we may care about the ordering of `enqueue` and `dequeue` operations while ignoring calls to `size` operation. In the remaining of this work, we use words `enq` and `deq` for the interests of space. Furthermore, we may allow `deq` operations to be executed out of order, but keep `enq` operations in order. In such case, we need a way to add ordering constraints on a subset of the methods of the shared object.

Let $Domain(o)$ be the set of all operations of a shared object o. Let $d \subset Domain(o)$ be a subset of operations. Let $Powerset(Domain(o))$ be the set of all subsets of $Domain(o)$. Let $D \subset Powerset(Domain(o))$ be a subset of the powerset.

Definition 3 (Quasi Linearization Factor)**.** *A* quasi-linearization factor *is a function $Q_O : D \rightarrow \mathbb{N}$, where D is a subset of the powerset and $\mathbb{N}$ is the set of natural numbers.*

Example 1. *For a bounded queue that stores a set X of non-zero data items, we have $Domain(\texttt{queue}) = \{enq.x, deq.x, deq.0 \mid x \in X\}$, where $enq.x$ denotes the `enqueue` operation for data x, $deq.x$ denotes the `dequeue` operation for data x, and $deq.0$ indicates that the queue is empty. We may* *define two subsets of $Domain(\texttt{queue})$:*

$$d_1 = \{enq.y \mid y \in Y\} \ ,$$

$$d_2 = \{deq.y \mid y \in Y\} \ .$$

Let $D = \{d_1, d_2\}$, where d_1 is the subset of `deq` *events and d_2 is the subset of* `enq` *events. The distance between σ and σ', after being projected to subsets d_1 and d_2, is defined as $\Delta(\sigma|_{d_1}, \sigma'|_{d_2})$. If we require that the* `enq` *calls follow the FIFO order and the* `deq` *calls be out-of-order by at most K steps, the quasi-linearization factor $Q_{\{\texttt{queue}\}} : D \to \mathbb{N}$ is defined as follows:*

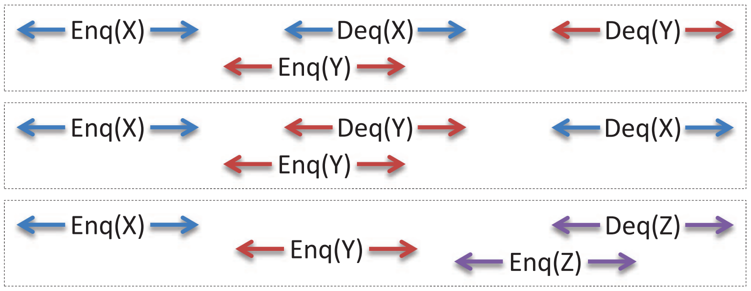
$$Q_{\{\texttt{queue}\}}(d_1) = 0 \ ,$$

$$Q_{\{\texttt{queue}\}}(d_2) = K \ .$$

Definition 4 (Quasi Linearizability). *Given a model $\mathcal{M} = (O = \{o_1, \ldots, o_k\}, init_O, \ P = \{p_1, \ldots, p_n\})$. Let H be the behavior of $\mathcal{M}$. $\mathcal{M}$ is quasi linearizable under the quasi factor $Q_O : D \to \mathbb{N}$ if for any history σ in H, there exists a sequential permutation π of σ such that*

- *for every op_1 and op_2 in σ, if $op_1 <_\sigma op_2$, then $op_1 <_\pi op_2$ (i.e., π respects the run-time ordering of operations), and*
- *for each object o_i ($1 \leq i \leq k$), there exists another sequential permutation π' of π such that*
  1. *$\pi'|_{o_i}$ is a legal sequential history (i.e., π' respects the sequential specification of the objects) and*
  2. *$\Delta((\pi|_{o_i})|_d, (\pi'|_{o_i})|_d) \leq Q_O(d)$ for all $d \in D$.*

This definition subsumes the definition for linearizability because, if the quasi factor is $Q_O(d) = 0$ for all $d \in D$, then the objects behaves as a standard linearizable data structure, e.g., a FIFO queue.



Only the first trace (at the top) is linearizable. The second trace is not linearizable, but is 1-quasi linearizable. The third trace is only 2-quasi linearizable.

Figure 2.1: Execution traces of a queue.

Example 2. *Consider the concurrent execution of a queue as shown in the Fig. 2.1. In the first part, it is clear that the execution is linearizable, because it is a valid permutation of the sequential history where `Enq(Y)` takes effect before `Deq(X)`. The second part is not linearizable, because the first dequeue operation is `Deq(Y)` but the first enqueue operation is `Enq(X)`. However, it is interesting to note that the second history is not far from a linearizable history, since swapping the order of the two dequeue events would make it linearizable. Therefore, flexibility is provided in dequeue events to allow them to be reordered. Similarly, for the third part, if the quasi factor is 0 (no out-of-order execution) or 1 (out-of-order by at most 1 step), then the history is not quasi linearizable. However, if the quasi factor is 2 (out-of-order by at most 2 steps), then the third history in Fig.2.1 is considered as quasi linearizable.*

2.3 Linearizability as Refinement

Linearizability is defined in terms of the invocations and responses of high-level operations. In a real concurrent program, the high-level operations are implemented by algorithms on concrete shared data structures, e.g., a linked list that implements a shared stack object [21]. Therefore, the execution of high-level operations may have complicated interleaving of low-level actions. Linearizability of a concrete concurrent algorithm requires that, despite low-level interleaving, the history of high-level invocation and response actions still has a sequential permutation that respects both the run-time ordering among operations and the sequential specification of the objects.

For verifying standard (but not quasi) linearizability, an existing method [15, 14] can be used to check whether a real concurrent algorithm (we refer as *implementation* in this work) refines the high-level linearizable requirement (we refer as *specification* in this work). In this case, the behaviors of the implementation and the specification are modeled as labeled transition systems (LTSs), and the refinement checking is accomplished by using explicit state model checking.

Definition 5 (Labeled Transition System). *A Labeled Transition System (LTS) is a tuple* $L = (S, init, Act, \rightarrow)$ *where* S *is a finite set of states;* $init \in S$ *is an initial state;* Act *is a finite set of actions; and* $\rightarrow \subseteq S \times Act \times S$ *is a labeled transition relation.*

For simplicity, we write $s \xrightarrow{\alpha} s'$ to denote $(s, \alpha, s') \in \rightarrow$. The set of enabled actions at s is $enabled(s) = \{\alpha \in Act \mid \exists s' \in S.\ s \xrightarrow{\alpha} s'\}$. A path π of L is a sequence of alternating states and actions, starting and ending with states $\pi = \langle s_0, \alpha_1, s_1, \alpha_2, \cdots \rangle$ such that $s_0 = init$ and $s_i \xrightarrow{\alpha_{i+1}} s_{i+1}$ for all i. If π is finite, then $|\pi|$ denotes the number of transitions in π. A path can also be infinite,

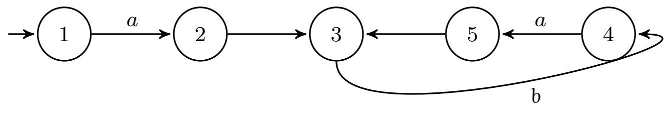


Figure 2.2: An LTS example

i.e., containing infinite number of actions. Since the number of states are finite, infinite paths are paths containing loops. The set of all possible paths for L is written as $paths(L)$.

A transition label can be either a visible action or an invisible one. Given an LTS L, the set of visible actions in L is denoted by vis_L and the set of invisible actions is denoted by $invis_L$. A τ-transition is a transition labeled with an invisible action. A state s' is *reachable* from state s if there exists a path that starts from s and ends with s', denoted by $s \overset{*}{\Rightarrow} s'$. The set of τ-successors is $\tau(s) = \{s' \in S \mid s \overset{\alpha}{\rightarrow} s' \wedge \alpha \in invis_L\}$. The set of states reachable from s by performing zero or more τ transitions, denoted as $\tau^*(s)$, can be obtained by repeatedly computing the τ-successors starting from s until a fixed point is reached. We write $s \overset{\tau*}{\rightarrow} s'$ iff s' is reachable from s via only τ-transitions, i.e., there exists a path $\langle s_0, \alpha_1, s_1, \alpha_2, \cdots, s_n\rangle$ such that $s_0 = s$, $s_n = s'$ and $s_i \overset{\alpha_{i+1}}{\rightarrow} s_{i+1} \wedge \alpha_{i+1} \in invis_L$ for all i . Given a path π, we can obtain a sequence of visible actions by omitting states and invisible actions. The sequence, denoted as $trace(\pi)$, is a trace of L. The set of all traces of L, is written as $traces(L)$ = $\{trace(\pi) \mid \pi \in paths(L)\}$.

LTSs can often be shown graphically, e.g., Fig. 2.2 shows an example LTS, where invisible transition labels are omitted for simplicity. We define the refinement relation between two LTSs, usually called trace refinement, as follows.

Definition 6 (Refinement). *Let L_1 and L_2 be two LTSs. L_1 refines L_2, written as $L_1 \sqsupseteq_T L_2$ iff*

$traces(L_1) \subseteq traces(L_2)$. □

In [15], we have shown that if L_{impl} is an implementation LTS and L_{spec} is the LTS of the linearizable specification, then L_{impl} is linearizable if and only if $L_{impl} \sqsupseteq_T L_{spec}$.

Given two labeled transition systems M_{spec} and M_{impl} representing specification and implementation LTS respectively, we say that M_{impl} refines M_{spec} if and only if the set of execution traces in M_{impl} is a subset of the execution traces in M_{spec}. So, the idea in refinement checking is to establish the (weak) simulation relationship between the specification model and the implementation model. The main approach is to perform the exhaustive search for the state space that is build of combined the specification-implementation. We compare every reachable state of the implementation with that of the specification reachable via same trace. If this condition suffices, we say that the M_{spec} refines M_{impl}, otherwise we get a counterexample which violates the check.

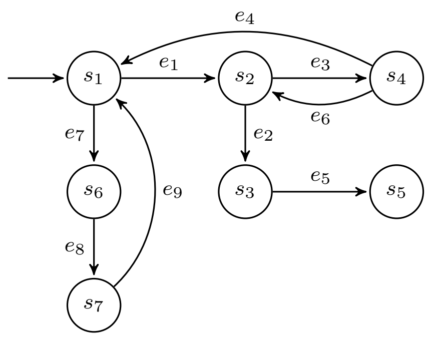


Figure 2.3: Specification model

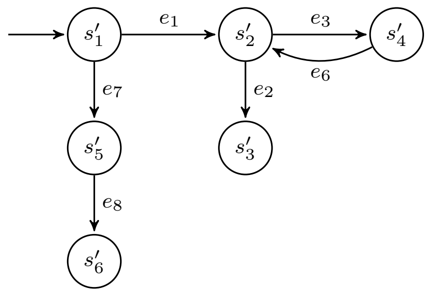


Figure 2.4: Implementation model

We start by presenting an example that illustrates the refinement checking. Fig. 2.3 shows the specification model and Fig. 2.4 shows the implementation model. Starting from the initial state s'_1 in the implementation model, we check the existence of a state with an enabled event e_1 from the s_1 in the specification model. As, the *Spec* LTS also has the same transition from state s_1, we continue the search by grabbing the next state. We can use Breadth-First Search (BFS) or Depth-First Search (DFS) to search the children. Considering a LTS with visible events e and invisible events τ, formally, we can grab the children of a state as follows:

1. if $impl \xrightarrow{\tau} impl'$, where τ is an internal event, then $spec' = spec$;

2. if $impl \xrightarrow{e} impl'$, where e is a method call event, then $spec \xrightarrow{e} spec'$;

We continue the check till all states of the implementation model is exhausted. In the given example, all the traces in implementation model shown in Fig. 2.4 are contained within the specification model shown in Fig. 2.3, meaning implementation refines specification.

2.4 Realization of a Model

We use a tool called PAT to model $spec$ and $impl$. It is a framework that supports the specification and verification of concurrent models. We model concurrent systems using a process algebra that is a variant of CSP. A LTS generated describes the behavior of the model. Basically CSP syntax is used in our model with the extension of shared variable. Table 2.1 shows the process definitions used in our model. Here CSP is extended such that shared variables are modeled alongside the processes to implement the different behaviors of the models.

Refinement checking is the method to verify if the behaviors of implementation follow the specification. In PAT, an assertion is specified to check the behaviors of two processes. Consider the two

Table 2.1: Process Definitions

Process	Definition
$P() = Stop$	A process that communicates nothing. It is also called the deadlock process.
$P() = Skip$	Terminates an event. $Skip$ is termination process.
$P() = event\{assignments\} \rightarrow P()$	Describes that $event$ is performed first and then it behaves like a process $P()$.
$P() = P_1(); P_2();$	It is the sequential composition of two processes P_1 and P_2 $P_1()$ is executed first and $P_2()$ is executed after that.
$P() = P_1() \vert\vert\vert P_2()$	It is the interleaved composition of the two processes $P_1()$ and $P_2()$.

processes $P()$ and $Q()$. The statement $\#assert\ \ P()\ \ refines\ \ Q()$ checks if there is a refinement relationship between these two processes $P()$ and $Q()$. The reachability analysis is conducted by exploring the product state transition of $P()$ and $Q()$, wherein the state that violates the refinement relationship is searched.

2.5 Related Work

In the literature, although there exists a large body of work on formally verifying linearizability in models of data structure implementations, none of them can verify quasi linearizability. For example, Liu et al. [15, 14] use a process algebra based tool to verify that an implementation model refines a specification model – the refinement relation implies linearizability. Vechev et al. [24] use the SPIN model checker to verify linearizability in a Promela model. Cerný et al. [5] use automated abstractions together with model checking to verify linearizability properties. There also exists some work on proving linearizability by constructing mechanical proofs, often with significant manual intervention (e.g., [22, 23]).

There are also runtime verification algorithms such as Line-Up [4], which can directly check the actual source code implementation but for violations on bounded executions and deterministic linearizability. However, quasi linearizable data structures are inherently nondeterministic. For example, the `deq` operation in a quasi queue implementation may choose to return any of the first k items in a queue. To the best of our knowledge, no existing method can directly verify quasi linearizability for execution traces of unbounded length.

Besides (quasi) linearizability, there also exist many other consistency conditions for concurrent computations, including sequential consistency [13], quiescent consistency [2], and eventual consistency [25]. Some of these consistency conditions in principle may be used for checking the correctness of data structure implementations, although so far, none of them is as widely used as (quasi) linearizability. These consistency conditions do not involve quantitative aspects of the properties. We believe that it is possible to extend our refinement algorithm to verify some of these properties. However, we leave it for future work.

Outside the domain of concurrent data structures, *serializability* and *atomicity* are two popular correctness properties for concurrent programs, especially at the application level. There exists a large body of work on both static and dynamic analysis for detecting violations of such properties (e.g., [8, 6] and [27, 7, 18, 26]). These existing methods are different from ours because they are checking different properties. Although atomicity and serializability are fairly general correctness conditions, they have been applied mostly to the correctness of shared memory accesses at the load/store instruction level. Linearizability, in contrast, defines correctness condition at the method call level. Furthermore, existing methods for checking atomicity and serializability do not deal with the quantitative aspects of the properties.

2.6 Summary

This chapter presented a brief overview of linearizability and quasi linearizability in relation to refinement. The idea of model checking by building the labeled transition systems and refinement checking was introduced. The chapter also included the related work on formal verification of linearizable data structures.

Chapter 3

Verifying Quasi Linearizability via Refinement Checking

This chapter presents our two new approaches for verifying quasi linearizability. We first explore two possible paths that allow us to verify quasi linearizability. First, we present the detailed description of a manual verification approach and show its limitations. This motivates our work in next chapter, which is extending the standard refinement checking algorithm to build an automated quasi linearizable checker.

3.1 Two Approaches of Verification

Our verification problem is defined as follows: Given an implementation model M_{impl}, a specification model M_{spec}, and a quasi factor Q_O, decide whether M_{impl} is quasi linearizable with respect to M_{spec} under the quasi factor Q_O.

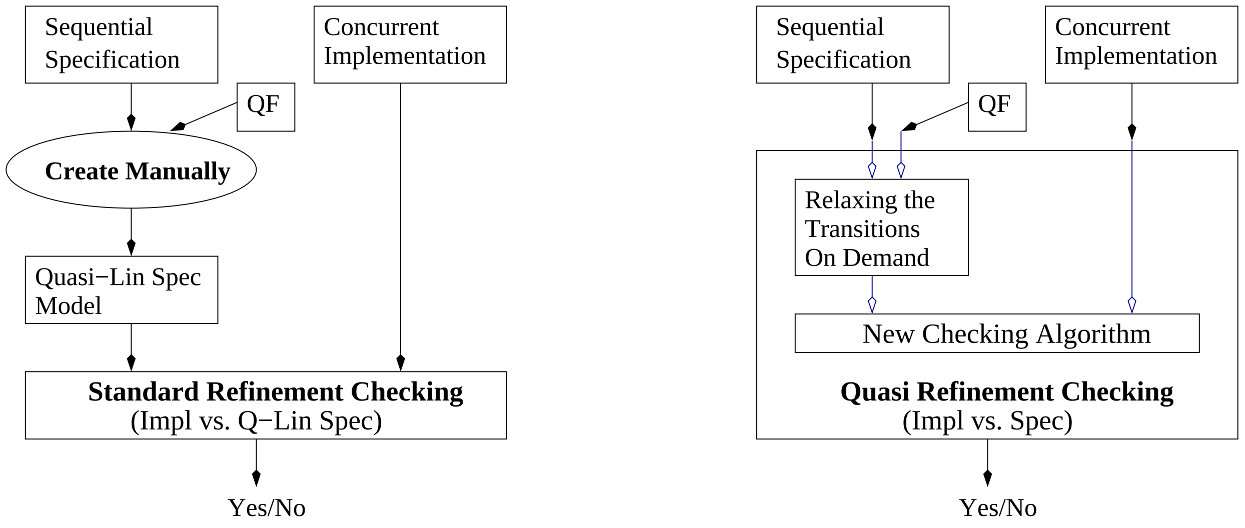


Figure 3.1: Verifying quasi linearizability: manual approach (left) and automated approach (right).

The straightforward approach for solving the problem is to leverage the procedure in Algorithm 3. However, since the procedure checks for standard refinement relation, not quasi refinement relation, the user has to manually construct a relaxed specification model, denoted M'_{spec}, based on the given specification model M_{spec} and the quasi factor Q_O. This so-called *manual approach* is illustrated by Fig. 3.1 (left). The relaxed specification model M'_{spec} must be able to produce all histories that can be produced by M_{spec}, as well as the new histories that are allowed under the relaxed consistency condition in Definition 4.

Unfortunately, there is no systematic method, or general guideline, on constructing such relaxed specification models. Each M'_{spec} may be different depending on the type of data structures to be checked. And there is significant amount of creativity required during the process, to make sure that the new specification model is both simple enough and permissive enough. For example, to verify that a K-segmented queue [1] is quasi linearizable, we can create a relaxed specification model whose `dequeue` method randomly removes one of the first K data items from the otherwise standard FIFO queue. This new model M'_{spec} will be more complex than M_{spec}, but can still be significantly simpler than the full-fledged implementation model M_{impl}, which requires the use of a complex segmented linked list.

Since the focus of this thesis is on designing a fully automated verification method, we shall briefly illustrate the manual approach in this chapter, and then focus on developing an automated approach in the subsequent chapter.

Our automated approach is shown in Fig. 3.1 (right). It is based on designing a new refinement checking algorithm that, in contrast to Algorithm 3, can directly check a *relaxed version* of the standard refinement relation between M_{impl} and M_{spec}. Therefore, the user does not need to manually construct the relaxed specification model M''_{spec}. Instead, inside the new refinement checking procedure, we systematically extend states and transitions of the specification model M_{spec} so that the new states and transitions as required by M'_{spec} are added on the fly. This would lead to the inclusion of a bounded degree of out-of-order execution on the relevant subset of operations as defined by the quasi factor Q_O. A main advantage of our new method is that the procedure is fully automated, thereby avoiding the user intervention, as well as the potential errors that may be introduced during the user's manual modeling process. Furthermore, by exploring the relaxed transitions on a *need-to* basis, rather than upfront as in the manual approach, we can reduce the number of states that need to be checked.

3.2 The Manual Verification Approach

In this section, we will briefly describe the manual approach and then focus on presenting the automated approach in the subsequent chapter. Although we do not intend to promote the manual approach – since it is labor-intensive and error prune – this section will illustrate the intuitions behind our fully automated verification method.

Given the specification model M_{spec} and the quasi factor Q_O, we show how to manually construct the relaxed specification model M'_{spec} in this section. We use the standard FIFO queue and two versions of quasi linearizable queues as examples. The construction needs to be tailored case by case for the different types of data structures.

Algorithm 1 Enqueue and Dequeue Pseudo code for Quasi-Specification model

```
1: Procedure enqAbs(tid) :=
2: if (HA < SIZE_ABS+1) then
3:     enq.tid.(HA+1) {HA= HA+1;}
4: end if
```

```
1: Procedure deqAbs(tid) :=
2: random_value = randomget();
3: deqvalue[tid] = random_value;
4: if (HA > deqvalue[tid]) then
5:     deq.tid.(HA-deqvalue[tid]){HA=HA
       -1;}
6: else if (HA == 0) then
7:     deq.tid.false
8: end if
```

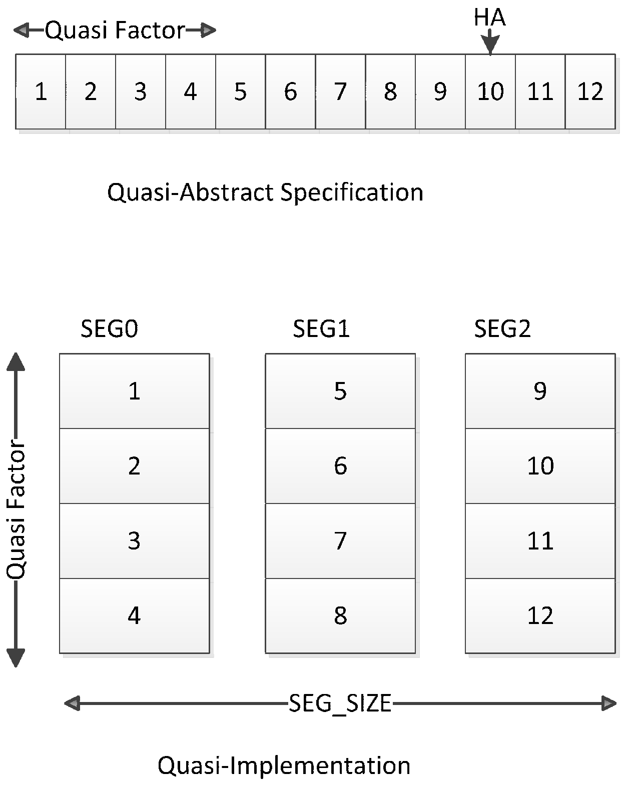


Figure 3.2: Implementations of a 4-quasi queue

3.2.1 Specification Model M_{spec}

The standard FIFO queue with a bounded capacity can be implemented by using a linked list, where `deq` operation removes a data item at one end of the list called the *head* node, and `enq` operation adds a data item at the other end of the list called the *tail* node. When the queue is full, `enq` does not have any impact. When the queue is empty, `deq` returns NULL. As an example, consider a sequence of four enqueue events `enq(1)`, `enq(2)`, `enq(3)`, `enq(4)`, the subsequent dequeue events would be `deq.1`, `deq.2`, `deq.3`, `deq.4`, which obey the FIFO semantics. This is illustrated by the first history `H1-a` in Fig. 3.3.

In the PAT model checking environment, the specification model M_{spec} is written in a process

```
H1-a      H1-b      H1-a      H1-b
-------   -------   -------   -------
enq(1)    enq(1)    enq(1)    enq(1)
enq(2)    enq(2)    enq(2)    enq(2)
enq(3)    enq(3)    enq(3)    enq(3)
enq(4)    enq(4)    enq(4)    enq(4)
deq()=1   deq()=1   deq()=2   deq()=2
deq()=2   deq()=2   deq()=1   deq()=1
deq()=3   deq()=4   deq()=3   deq()=4
deq()=4   deq()=3   deq()=4   deq()=3
-------   -------   -------   -------
```

Here, `deq` can be out-of-order by 1. The first `deq` randomly returns a value from the set $\{1, 2\}$ and the second `deq` returns the remaining one. Then the third `deq` randomly returns a value from the set $\{3, 4\}$ and the forth `deq` returns the remaining one.

Figure 3.3: Valid histories of a *1-quasi linearizable* queue

algebra language, named CSP# [19]. Algorithm 1 shows the detailed explaination for creating an abstract model in the PAT tool environment.

3.2.2 Implementation Model M_{impl}:

The bounded quasi linearizable queue can be implemented by using a segmented linked list. This is the original algorithm proposed by Afek *et al.* [1]. A segmented linked list is a linked list where each list node can hold K data items, as opposed to a single data item in the standard linked list. As shown in Fig. 3.2 (lower half), these K data items form a *segment*, in which the data slots are numbered as 1, 2, . . ., K. In general, the segment size needs to be set to $(QF + 1)$, where QF is the maximum number of out-of-order execution steps. The example in Fig. 3.2 has the quasi factor set to 3, meaning that a `deq` operation can be executed out of order by at most 3 steps. Consequently, the size of each segment is set to (3+1)=4. Since $Q_{\{\texttt{queue}\}}(D_{\texttt{enq}}) = 0$, meaning that the `enq`

Algorithm 2 Enqueue and Dequeue Pseudocode for Quasi-implementation model

```
Enq(tid) :=
cur_node[tid] = 0; cur_seg[tid] = 0
while $cur\_seg[tid] < SIZE\_SEG$ do
    while $cur\_node[tid] < QF$ do
        if $flag\_imp[cur\_seg[tid]][cur\_node[tid]] == 0$ then
            flag_imp[cur_seg[tid]][cur_node[tid]] = 1
            item_count++
            enq.tid.item_count
            break
        else
            cur_node++
        end if
    end while
    cur_node = 0
    cur_seg++
end while

Deq(tid) :=
cur_seg[tid] = 0; cur_node[tid] = 0
while $cur\_seg < SIZE\_SEG$ do
    while $cur\_node < QF$ do
        cur_node[tid] = perm1[cur_node[tid]]
        if $flag\_imp[cur\_seg[tid]][cur\_node[tid]] == 1$ then
            flag_imp[cur_seg[tid]][cur_node[tid]]= 0
            item_count- -
            deq.tid.(item_count+1)
            break
        else
            cur_node++
        end if
    end while
    cur_seg++
    cur_node = 0
end while
if $item\_count == 0$ then
    deq.tid.false
end if
```

operations cannot be reordered, the data items are enqueued regularly in the empty slots of one segment, before the *head* points to the next segment. But for `deq` operations, we randomly remove one existing data item from the current segment. Algorithm 2 shows the detailed explaination for creating an implementation model in the PAT tool environment.

3.2.3 Relaxed Specification Model M'_{spec}:

Not all execution traces of M_{impl} are traces of M_{spec}. In Fig. 3.3, histories other than `H1-a` are not linearizable. However, they are all quasi linearizable under the quasi factor 1. They may be produced by a segmented queue where the segment size is (1+1)=2. To verify that M_{impl} is quasi linearizable, we construct a new model M'_{spec}, which includes not only all histories of M_{spec} but also the histories that are allowed only under the relaxed consistency condition. In this example, we choose to construct the new model by slightly modifying the standard FIFO queue. This is illustrated in Fig. 3.2 (upper half), where the first K data items are grouped into a cluster. Within the cluster, the `deq` operation may remove any of the k data items based on randomization. Only after the first k data items in the cluster are retrieved, will the `deq` move to the next k data items (a new cluster). The external behavior of this model is expected to match that of the segmented queue in M_{impl}: both are *1-quasi linearizable*.

Algorithm 3 Standard Refinement Checking

```
1: Procedure Check-Refinement($(impl, spec)$)
2: checked := $\emptyset$
3: pending.push($(init_{impl}, init_{spec})$)
4: while pending $\neq \emptyset$ do
5:     $(impl, spec)$ := pending.pop()
6:     if $enabled(impl) \not\subseteq enabled(spec)$ then
7:         return false
8:     end if
9:     checked := checked $\cup\{(impl, spec)\}$
10:    for all $(impl', spec') \in next(impl, spec)$ do
11:        if $(impl', spec') \notin$ checked then
12:            pending.push($(impl', spec')$)
13:        end if
14:    end for
15: end while
16: return true
```

3.2.4 Checking Refinement Relation:

Once M'_{spec} is available, checking whether M_{impl} refines M'_{spec} is straightforward by using Algorithm 3. Algorithm 3 shows the pseudo code of the refinement checking procedure in [15, 14]. Assume that L_{impl} refines M_{spec}, then for each reachable transition in M_{impl}, denoted as $impl \xrightarrow{e} impl'$, there must exist a reachable transition in L_{spec}, denoted as $spec \xrightarrow{e} spec'$. Therefore, the procedure starts with the pair of initial states of the two models, and repeatedly checks whether they have matching successor states. If the answer is no, the check at Lines 6-8 would fail, meaning that L_{impl} is not linearizable. Otherwise, for each pair of immediate successor states $(impl', spec')$, we add the pair to the *pending* list. The entire procedure continues until either (1) a non-matching transition in L_{impl} is found at Lines 6-8, or (2) all pairs of reachable states are checked, in which case L_{impl} is proved to be linearizable.

In Algorithm 3, the subroutine $next(impl, spec)$ is crucially important. It takes the current states of L_{impl} and L_{spec} as input, and returns a set of state pairs of the form $(impl', spec')$. Here each pair $(impl', spec')$ is one of the immediate successor state pairs of $(impl, spec)$. They are defined as follows:

1. if $impl \xrightarrow{\tau} impl'$, where τ is an internal event, then let $spec' = spec$;
2. if $impl \xrightarrow{e} impl'$, where e is a method call event, then $spec \xrightarrow{e} spec'$;

We have assumed, without loss of generality, that the specification model L_{spec} is deterministic. If the original specification model is nondeterministic, we can always apply standard *subset construction* (of DFAs) to make it deterministic.

3.3 Experimental Results

For the segmented queue implementation [1], we have manually constructed M'_{spec} and checked the refinement relation in the PAT model checking environment. Our experimental results are summarized in Table 3.1. Column 1 shows the different quasi factors. Column 2 shows the number of segments – the capacity of the queue is $(QF + 1) \times Seg$. Column 3 shows the refinement checking time in seconds. Column 4 shows the total number of visited states during refinement checking. Column 5 shows the total number of state transitions activated during refinement checking. The experiments are conducted on a computer with an Intel Core-i7, 2.5 GHz processor and 8 GB RAM running Ubuntu 10.04.

Table 3.1: Experimental results for standard refinement checking.

Quasi Factor	#. Segment	Verification Time (s)	#. Visited State	#. Transition
1	1	0.1	423	778
1	2	0.1	2310	4458
1	3	0.1	8002	15213
1	4	0.4	22327	41660
1	5	0.9	55173	101443
1	6	2.0	126547	230259
1	10	55.9	2488052	4421583
1	15	MOut	-	-
2	1	0.6	26605	58281
2	2	12.6	456397	970960
2	3	130.7	4484213	8742485
2	4	MOut	-	-
3	1	8.8	284484	638684
3	2	MOut	-	-
4	1	124.4	3432702	7906856
4	2	MOut	-	-

MOut means memory-out.

The experimental results in Table 3.1 show an exponential increase in the verification time when we increase the size of the queue or the quasi factor. For size = 1, verification completes as soon as 0.1 seconds. When it is increased to 2, it takes around 0.6 seconds to check if the implementation model refines specification or not. This is inevitable since the size of the state space grows exponentially as the size increases. Furthermore, as the quasi factor increases, the verification time is also significantly increased. We can see that for quasi factor as small as 1, it takes only 0.1 seconds to complete the verification. But, when quasi factor is increased to 2 and 3 subsequently, there is tremendous increase in the visited states count and hence the verification time. Fig. 3.4 shows the relation of quasi factor and the verification time. The horizontal axis represents the quasi factor. Primary vertical axis represents the verification time for the corresponding quasi factor and the segment size. And, the secondary vertical axis represents the segment size of the queue.

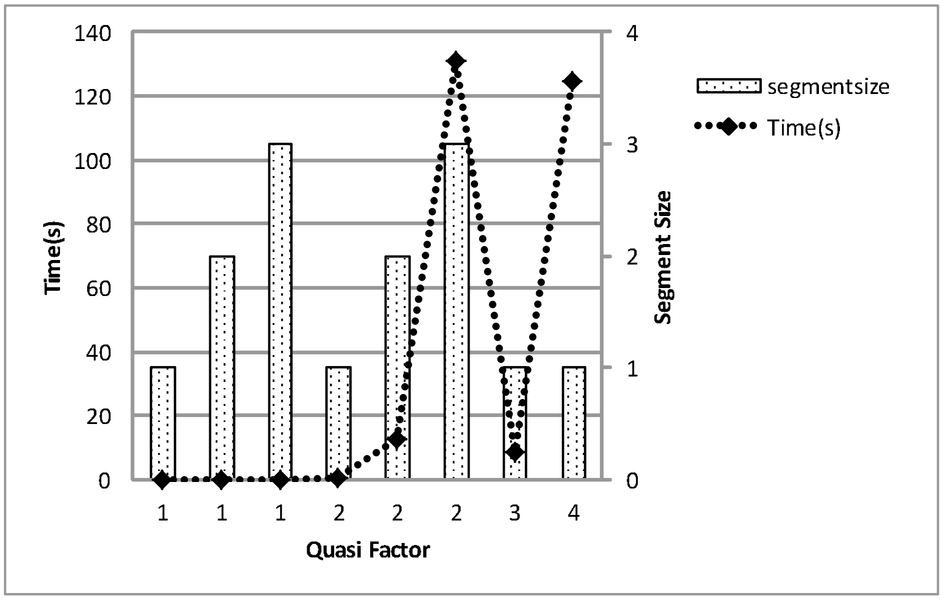


Figure 3.4: Graph showing the linearizability checking time for different QFs and segment sizes of a queue

3.4 Limitations of the Manual Approach

The proposed method is limited by the availability of the specification model in order to compare with the implementation models. One needs to know the possible traces for a model in order to check the correctness of the implementation model. So, this method requires the user to manually construct M'_{spec}, which is a severe limitation because it is often labor intensive and error prone. For example, consider the seemingly simple random dequeued model in Fig. 3.2. A subtle error would be introduced if we do not use the *cluster* to restrict the set of data items that can be removed by `deq` operation. Assume that `deq` always returns one of the first k data items in the current queue. Although it appears to be correct, such implementation will not be k-quasi linearizable, because it is possible for some data item to be over-taken indefinitely. For example, if every time `deq` chooses *the second data item in the list*, we will have the following `deq` sequence: `deq.2`, `deq.3`, `deq.4`, $\ldots$, `deq.1`, where the dequeue of value 1 can be delayed by an arbitrarily long time. This is no longer a *1-quasi linearizable* queue. In other words, if the user construct M'_{spec} incorrectly, the verification result becomes invalid.

Therefore, we need to design a fully automated method to directly verify quasi linearizability of M_{impl} against M_{spec} under the given quasi factor QF.

3.5 Summary

In this chapter, we proposed a model checking approach to verify the quasi linearizability of a concurrent data structure. Since, this technique is labor intensive and error prone, we proceed with the fully automated approach for checking the *Quasi Linearizability* in the next chapter.

Chapter 4

New Algorithm for Checking the Quasi Refinement Relation

In this chapter, we present our automated method for verifying *Quasi Linearizability*. We shall start with the standard refinement checking procedure in Algorithm 3 and extend it to directly check a relaxed version of the refinement relation between M_{impl} and M_{spec} under a given quasi factor. The idea is to establish the simulation relationship from specification to implementation while allowing relaxation of the specification. This not only reduces the chances of committing error during specification construction phase, but also automates the tool which can verify the correctness of concurrent data structures.

4.1 Motivation

The manual verification approach presented in Chapter 3 not only can check the validity of the concurrent data structures, but also can generate the counterexample if the implementation model does not refine the specification. However, this is obtained at the cost of extensive manual effort used to model the quasi linearizable specification, which should contain all the possible traces that a relaxed linearizable implementation could have. Also, the possibility of introducing errors while writing the specification model is high. Therefore, we propose a new and fully automated method which relaxes the specification of a model first and then conduct refinement checking to ascertain the correctness of the implementation model.

4.2 Linearizability Checking via Quasi Refinement

The new procedure, shown in Algorithm 4, is different from Algorithm 3 as follows:

1. We customize *pending* to make the state exploration follow a breadth-first search (BFS). In Algorithm 3, it can be either BFS or DFS based on whether *pending* is a queue or stack.

2. We replace *enabled(spec)* with *enabled_relaxed(spec,QF)*. It will return not only the events enabled at current $spec$ state in M_{spec}, but also the additional events allowed under the relaxed consistency condition.

3. We replace *next(impl,spec)* with *next_relaxed(impl,spec,QF)*. It will return not only the successor state pairs in the original models, but also the additional pairs allowed under the

Algorithm 4 Quasi Refinement Checking

```
1: Procedure Check-Quasi-Refinement$(impl, spec, QF)$
2: checked := $\emptyset$
3: pending.enqueue$((init_{impl}, init_{spec}))$
4: while pending $\neq \emptyset$ do
5:    $(impl, spec)$ := pending.dequeue()
6:    if $enabled(impl) \not\subseteq enabled\_relaxed(spec, QF)$ then
7:       return false
8:    end if
9:    checked := checked $\cup \{(impl, spec)\}$
10:   for all $(impl', spec') \in next\_relaxed(impl, spec, QF)$ do
11:      if $(impl', spec') \notin$ checked then
12:         pending.enqueue$((impl', spec'))$
13:      end if
14:   end for
15: end while
16: return true
```

relaxed consistency condition.

Conceptually, it is equivalent to first constructing a relaxed specification model M'_{spec} from (M_{spec}, QF) and then computing the *enabled(spec)* and *next(impl,spec)* on this new model. However, in this case, we are constructing M'_{spec} automatically, without the user's intervention. Furthermore, the additional states and edges that need to be added to M'_{spec} are processed incrementally, on a *need-to* basis.

At the high level, the new procedure performs a BFS exploration for the state pair $(impl, spec)$, where *impl* is the state of implementation and *spec* is a state of specification. The initial implementation and specification events are enqueued into *pending* and each time we go through the while-loop, we dequeue from *pending* a state pair, and check if all events enabled at state $impl$ match with some events enabled at state $spec$ under the relaxed consistency condition (line 6).

If there is any mismatch, the check fails and we can return a counterexample showing how the violation happens. Otherwise, we continue until *pending* is empty. Lines 10-14 explore the new successor state pairs, by invoking *next_relaxed* and add to *pending* if they have not been checked.

Subroutine enabled_relaxed(spec,QF): It takes the current state $spec$ of model M_{spec}, along with the quasi factor QF, and generates all events that are enabled at state $spec$.

Consider the graph in Fig. 4.1 as M_{spec}. Without relaxation, $enabled(s_1)$=$\{e_1\}$. This is equivalent to $enabled\_relaxed(s_1, 0)$. However, when $QF = 1$, according to the dotted edges in Fig. 4.2, the set $enabled\_relaxed(s_1, 1)$=$\{e_1, e_2, e_3\}$.

The reason why e_2 and e_3 become enabled is as follows: before relaxation, starting at state s_1, there are two length-3 $(2QF + 1)$ event sequences $\sigma_1 = e_1, e_2, e_5$ and $\sigma_2 = e_1, e_3, e_4$. When $QF = 1$, it means an event can be out-of-order by at most 1 step. Therefore, the possilbe valid permutations of σ_1 is $\pi_1 = e_2, e_1, e_5$ and $\pi_2 = e_1, e_5, e_2$, and the possible valid permutations of σ_2 is $\pi_3 = e_3, e_1, e_4$ and $\pi_4 = e_1, e_4, e_3$ for $QF = 1$. In other words, at state s_1, events e_2, e_3 can also be executed. We will discuss the generation of valid permutation sequences in Section 4.3.

Subroutine next_relaxed$(impl, spec, QF)$**:** It takes the current state $impl$ of M_{impl} and the current state $spec$ of M_{spec} as input, and returns a set of state pairs of the form $(impl', spec')$. Similar to the definition of $next(impl, spec)$ in Section **??**, we define each pair $(impl', spec')$ as follows:

1. if $impl \xrightarrow{\tau} impl'$, where τ is an internal event, then let $spec' = spec$;

2. if $impl \xrightarrow{e} impl'$, where e is a method call event, then $spec \xrightarrow{e} spec'$ where event $e \in enabled\_relaxed(spec, QF)$ is enabled at $spec$ after relaxation.

For example, when $spec = s_1$ in Fig. 4.1, and the quasi factor is set to 1 – meaning that the event at state s_1 can be out-of-order by at most one step – the procedure *next_relaxed(impl,*s_1*, 1)* would return not only $(impl', s_2)$, but also $(impl', s_6)$ and $(impl', s_9)$, as indicated by the dotted edges in Fig. 4.2. The detailed algorithm for generation of the relaxed next states in specification is described in Section 4.3.

4.3 Generation of Relaxed Specification

In this subsection, we show how to relax the specification M_{spec} by adding new states and transitions – those that are allowed under the condition of quasi linearizability – to form a new specification model. Notice that we accomplish this automatically, and incrementally, on a *need-to* basis.

For each state $spec$ in M_{spec}, we compute all the event sequences starting at $spec$ with the length $(2QF + 1)$. These event sequences can be computed by using a simple graph traversal algorithm, e.g., a breadth first search.

Fig. 4.1 shows an example for the computation of these event sequences. The specification model M_{spec} has the following set of states $\{s_1, s_2, s_3, s_4, s_5\}$. Suppose that the current state is s_1 (in $step$ 0), then the current frontier state set is $\{s_1\}$, and the current event sequence is $\langle s_1 \rangle$. The

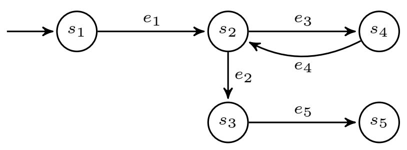


Figure 4.1: Specification model before adding relaxed transitions for state s_1.

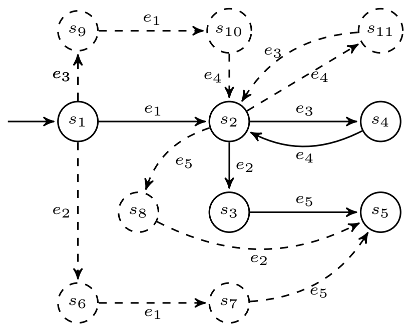


Figure 4.2: Specification model after adding relaxed edges for state s_1 and quasi factor 1.

Table 4.1: Specification sequence generation at state s_1

BFS Steps	(Frontier)	EventSequences
step 0	$\{s_1\}$	$\langle s_1 \rangle$
step 1	$\{s_2\}$	$\langle s_1 \xrightarrow{e_1} s_2 \rangle$
step 2	$\{s_3, s_4\}$	$\langle s_1 \xrightarrow{e_1} s_2 \xrightarrow{e_2} s_3 \rangle \langle s_1 \xrightarrow{e_1} s_2 \xrightarrow{e_3} s_4 \rangle$
step 3	$\{s_5, s_2\}$	$\langle s_1 \xrightarrow{e_1} s_2 \xrightarrow{e_2} s_3 \xrightarrow{e_5} s_5 \rangle \langle s_1 \xrightarrow{e_1} s_2 \xrightarrow{e_3} s_4 \xrightarrow{e_4} s_2 \rangle$

results of each BFS step are shown in Table 4.1. In $step$ 1, the frontier state set is $\{s_2\}$, and the event sequence becomes $\langle s_1 \xrightarrow{e_1} s_2 \rangle$. In $step$ 2, the frontier state set is $\{s_3, s_4\}$, and the event sequence is split into two sequences. One is $\langle s_1 \xrightarrow{e_1} s_2 \xrightarrow{e_2} s_3 \rangle$ and the other is $\langle s_1 \xrightarrow{e_1} s_2 \xrightarrow{e_3} s_4 \rangle$. The traversal continues until the BFS depth reaches ($2QF + 1$).

After completing the ($2QF + 1$) steps of BFS starting at state $spec$, as above, we have to generate possible valid permutations first and then we will be able to evaluate the two subroutines: $enabled\_relaxed(spec, QF)$ and $next\_relaxed(impl, spec, QF)$.

We transform the original specification model in Fig. 4.1 to the relaxed specification model in Fig. 4.2 for $QF = 1$. The dotted states and edges are newly added to reflect the relaxation. More specifically, for $QF = 1$, we will reach $(2QF + 1) = 3$ steps during the BFS. At $step$ 3, there are two existing sequences $\{e_1, e_2, e_5\}$ and $\{e_1, e_3, e_4\}$. For each existing sequence, we compute all possible valid permutation sequences. In this case, the valid permutation sequences are $\{e_2, e_1, e_5\}$,$\{e_1, e_5, e_2\}$ and $\{e_3, e_1, e_6\}$, $\{e_1, e_3, e_6\}$. For each newly generated permutation sequence, we add new edges and states to the specification model. From an initial state s_1, if we follow the new permutation $\{e_2, e_1, e_5\}$, as shown in Fig. 4.2, the transition e_2 will lead to newly formed pseudo state s_6, the transition e_1 will lead to s_7 from state s_6 and from this state it is reconnected back to the original state s_5 via transition e_5. Similarly, if we follow the new permutation $\{e_3, e_1, e_4\}$, the transition e_3 will lead to newly formed pseudo state s_9, the transition e_1 will lead to s_{10} from state s_9 and from this state it is reconnected back to state s_2 via transition e_4. We continue this process of state expansion for all the valid permutation sequences. This relaxation process needs to be conducted by using every existing state of M_{spec} as the starting point (for BFS up to $2QF+1$ steps) and then adding the new states and edges. Note that this process is conducted on the fly.

Algorithm 5 explains the high level pseudo-code for expanding the state space for the current specification state under the check. Let $SEQ = \{seq_1, seq_2, ..., seq_k\}$ be the sequences which are reachable from the state s_0 in M_{spec} such that each sequence has less than or equal to $2QF + 1$ events. Each sequence $seq \in SEQ$ calls a *genValidPermut(seq,QF)* (line 4) to generate all the possible valid permutation paths for that trace. A new state is formed with a new transition for each event in the permuted sequences, hence allowing the relaxed refinement checking of the implementation trace.

The valid permutations for a given sequence is generated using an Algorithm 6 which is based on the cost associated with the event. Initially, for each events e_i where $1 \leq i < n$ associated with the

Algorithm 5 Pseudo-code for Expanding Specification Under Check

```
1: Let $s_0$ be a specification state and QF be the quasi factor
2: Let $SEQ = \{seq_1, seq_2, seq_3, \cdots, seq_k\}$ be the set of all possible event sequences reachable
   from $s_0$ in $M_{spec}$ such that for $1 \leq i \leq k$, each $seq_i$ has less than or equal to 2QF $+ 1$ relaxed
   events
3: for all seq in SEQ do
4:    PERMUT_VALID = $genValidPermut(seq, QF)$
5:    for all perm in PERMUT_VALID do
6:       Let perm = $\langle e_1, e_2, \cdots, e_n \rangle$
7:       Let $s_n$ be the specification state reached from $s_0$ via seq
8:       if perm is not equal to seq then
9:          for all $e_i$ where $1 \leq i < n$ do
10:            Create a new state $s_i$ and a new transition from $s_{i-1}$ to $s_i$ via event $e_i$
11:         end for
12:         Create a new transition from $s_{n-1}$ to $s_n$ via $e_n$
13:      end if
14:   end for
15: end for
```

seq, the cost is initialized to *QF* (line 2). We generate all possible permutations and update cost with respect to the relative ordering of the events for each reshuffled sequences. This cost attribute of an event stores the information on how many more steps an event may be postponed. Each time an event is postponed, the cost associated with this event is decremented by 1. On the contrary, the event can also be chosen upto QF steps ahead and for each step, the cost is increased by 1. So, the cost attribute of the event that is allowed for relaxation is $2QF \leq$ cost ≤ 0. We check the validity of each of these sequences using this cost attribute (line 8). Finally, only the valid permutations are appended in *PERMUT_VALID* after each check and once the check is completed for all permuted sequences, the function returns the valid traces.

Consider the event sequence $\{e_1, e_2, e_5\}$ from state s_1 be *seq* as shown in Fig. 4.1. If $QF = 1$, the cost for each of these events is initialized to 1. We generate all possible permutations by reshuffling

Algorithm 6 $genValidPermut(seq, QF)$

```
1:  PERMUT_VALID := ∅
2:  Initialize cost associated with each event in seq to QF
3:  Generate possible permutations PERMUT_SEQ and update cost
4:  for all p in PERMUT_SEQ do
5:      isValid = true
6:      Let p = $\langle e_1, e_2, \cdots, e_n \rangle$
7:      for all $e_i$ where $1 \leq i < n$ do
8:          if $e_i.cost \geq 2QF \vee e_i.cost \leq 0$ then
9:              isValid = false
10:             break
11:         end if
12:     end for
13:     if isValid then
14:         PERMUT_VALID = PERMUT_VALID $\bigcup$ p
15:     end if
16: end for
17: return PERMUT_VALID
```

the events and updating the cost based on the relative positioning of the event with respect to the initial sequence. There are as many as 6 possible permutations including the original sequence in this case. If we consider reordering be the sequence $\{e_2, e_1, e_5\}$, then the cost associated with event e_2 is 2 as it is chosen one step earlier. For the event e_1, it is postponed for one step meaning its cost is decreased by 1 which makes the cost associated with it be 0. Event e_3 is not reordered and hence its cost is unchanged and is 1. This sequence is valid because cost associated with each of the events in this sequence lies within the allowable range. Similarly, if we consider another permuted sequence $\{e_3, e_1, e_2\}$, then the cost associated with each of these events is $\{3, 0, 0\}$ which exceeds the allowable range. So, this permutation sequence is not valid. We do this for all the permuted sequences to generate the valid traces.

4.4 Experimental Results

We have implemented and evaluated the quasi linearizability checking method in the PAT verification framework [20]. Our new algorithm can directly check a relaxed version of the refinement relation. This new algorithm subsumes the standard refinement checking procedure that has already been implemented in PAT. In particular, when $QF = 0$, our new procedure degenerates to the standard refinement checking procedure. When $QF > 0$, our new procedure has the added capability of checking for the quantitatively relaxed refinement relation. Our algorithm can directly handle the implementation model M_{impl}, the standard (not quasi) specification model M_{spec}, and the quasi factor QF, thereby completely avoiding the user's intervention.

Table 4.2: Statistics of the benchmark examples

Class	Description	Linearizable	Quasi Lin.
Quasi Queue (3)	Segmented linked list implementation (size=3)	No	Yes
Quasi Queue (6)	Segmented linked list implementation (size=6)	No	Yes
Quasi Queue (9)	Segmented linked list implementation (size=9)	No	Yes
Quasi Queue (4)	Segmented linked list implementation (size=4)	No	Yes
Quasi Queue (8)	Segmented linked list implementation (size=8)	No	Yes
Queue buggy1	Segmented queue with a bug (Dequeue on the empty queue may erroneously change current segment)	No	No
Queue buggy2	Segmented queue with a bug (Dequeue may get value from a wrong segment)	No	No
Lin. Queue	A linearizable (hence quasi) implementation	Yes	Yes
Q. Priority Queue (6)	Segmented linked list implementation (size=6)	No	Yes
Q. Priority Queue (9)	Segmented linked list implementation (size=9)	No	Yes
Q. Priority Queue (4)	Segmented linked list implementation (size=4)	No	Yes
Priority Queue buggy	Segmented priority queue (Dequeue on the empty priority queue may change current segment)	No	No
Lin. Stack	A linearizable (hence quasi) implementation	Yes	Yes

We have evaluated our new algorithm on a set of models of standard and quasi linearizable concurrent data structures [1, 12, 17], including queues, stacks, quasi queues, quasi stacks, and quasi priority queues. For each data structure, there can be several variants, each of which has a slightly different implementation. In addition to the implementations that are known to be linearizable

Table 4.3: Results for checking quasi linearizability with 2 threads

Class	QF	Verification Time (s)	Number of Visited States	Number of Visited Transitions
Quasi Queue (3)	2	7.2	126,810	248,122
Quasi Queue (6)	2	21.2	237,760	468,461
Quasi Queue (9)	2	114.5	1,741,921	3,424,280
Quasi Queue (4)	3	131.6	442,558	869,129
Quasi Queue (8)	3	1517.1	1,986,924	3,754,489
Queue buggy1	2	0.4	1,204	809
Queue buggy2	2	0.1	345	345
Lin. Queue	2	5.5	240,583	121,548
Q. Priority Queue (6)	2	34.3	472,981	918,530
Q. Priority Queue (9)	2	198.4	1,478,045	2,905,016
Q. Priority Queue (4)	3	343.1	1,408,763	2,566,427
Priority Queue buggy	2	5.4	894	894
Lin. Stack	2	0.2	2,690	6,896

and quasi linearizable, we also have versions which initially were thought to be correct, but were subsequently proved to be buggy by our verification tool. The characteristics of all benchmark examples are shown in Table 4.2. The first two columns list the name of the concurrent data structures and a short description of the implementation. The next two columns show whether the implementation is linearizable and quasi linearizable.

Table 4.3 shows the results of the experiments. The experiments are conducted on a computer with an Intel Core-i7, 2.5 GHz processor and 8 GB RAM running Windows 7. The first column shows the statistics of the test program, including the name and the size of benchmark. The second column is the quasi factor showing the relaxation bound allowed for the model. The next three columns show the runtime performance, consisting of the verification time in seconds, the total number of visited states, and the total number of transitions made. The number of states and the running time for each of the models increase with the data size.

For 3 segmented quasi queue with quasi factor 2, the verification completes in 7.2 seconds. It is much faster than the first approach presented in Section 4, where the same setting requires 130.7 seconds for the verification. Subsequently, as the size increases, the time to verify the quasi queue increases. For queue with size 6 and 9, verification is completed in 21.2 seconds and 114.5 seconds, respectively. As the quasi factor is increased to 3, the verification time for quasi queue with size 4 and 8 is increased to 131.6 seconds 1517.1 seconds respectively, which is much higher in comparison to the time for quasi factor 2. This is basically because of the significant increment in state expansion for the higher quasi factor. For the priority queues where enqueue and dequeue operations are performed based on the priority, the verification time is higher than the regular quasi queue. Also, it is important to note that the counterexample is produced with exploration of only part of the state space for the buggy models. The verification time is much faster for the buggy queue, which shows that our approach is effective if the quasi linearizability is not satisfied. In all test cases, our method was able to correctly verify quasi linearizability or detect the violations.

4.5 Conclusion

In this chapter, we presented a novel approach to automate the quasi linearizability refinement checking algorithm presented in Chapter 3. The specification model is relaxed on the fly by using our new refinement checking algorithm. This makes the method more robust and less error prone as the specification is relaxed automatically as opposed to manually by the user.

Chapter 5

Conclusions and the Future Work

In this thesis, we have presented a new method for formally verifying quasi linearizability of the implementation models of concurrent data structures. We have explored two approaches, one of which is based on manual construction of the relaxed specification model, whereas the other is fully automated, and is based on checking a relaxed version of the refinement relation between the implementation model and the specification model. The key idea of this work is to construct the quasi linearizable specification manually initially to analyze the bulkiness of the work and the cases of committing errors in modeling those specification models. This provides us a framework to extend the original refinement algorithm in order to verify the relaxed implementation model by automating the checker.

Experimental results have showed that our approach is not only able to verify the linearizable models, but also the relaxed linearizable models. The concurrent models with bugs have also been identified correctly in all cases. The counterexample traces generated for the buggy models clearly can provide us valuable information on why the models being buggy. Based on such information, we have been able to fix the bugs in our original models to get the correct versions of the implementation models.

The main objective of this thesis work is to develop an automated software tool that can formally verify a relaxed version of the linearizability property in the models of concurrent data structures. The main bottleneck of our approach is state space expansion, meaning the state space increases exponentially for higher number of processes as well as size of model. Also, it is limited by the fact that we can only verify the model with a bounded size to ensure that it has a finite state space. We believe that the automated refinement checking algorithm can be further optimized to improve the performance. For future work, we plan to incorporate advanced state space reduction techniques such as symmetry reduction and partial order reduction [14].

Bibliography

[1] Y. Afek, G. Korland, and E. Yanovsky. Quasi-Linearizability: Relaxed consistency for improved concurrency. In *International Conference on Principles of Distributed Systems*, pages 395–410, 2010.

[2] J. Aspnes, M. Herlihy, and N. Shavit. Counting networks. *J. ACM*, 41(5):1020–1048, 1994.

[3] H. Attiya and J. Welch. *Distributed Computing: Fundamentals, Simulations, and Advanced Topics*. John Wiley & Sons, Inc., Publication, 2nd edition, 2004.

[4] S. Burckhardt, C. Dern, M. Musuvathi, and R. Tan. Line-up: a complete and automatic linearizability checker. In *ACM SIGPLAN Conference on Programming Language Design and Implementation*, pages 330–340, 2010.

[5] P. Cerný, A. Radhakrishna, D. Zufferey, S. Chaudhuri, and R. Alur. Model checking of linearizability of concurrent list implementations. In *International Conference on Computer Aided Verification*, pages 465–479, 2010.

[6] A. Farzan and P. Madhusudan. Causal atomicity. In *International Conference on Computer Aided Verification*, pages 315–328, 2006.

[7] A. Farzan and P. Madhusudan. Monitoring atomicity in concurrent programs. In *International Conference on Computer Aided Verification*, pages 52–65, 2008.

[8] C. Flanagan and S. Qadeer. A type and effect system for atomicity. In *ACM SIGPLAN Conference on Programming Language Design and Implementation*, pages 338–349, 2003.

[9] T. A. Henzinger, A. Sezgin, C. M. Kirsch, H. Payer, and A. Sokolova. Quantitative relaxation of concurrent data structures. In *ACM SIGACT-SIGPLAN Symposium on Principles of Programming Languages*, 2013.

[10] M. Herlihy and J. M. Wing. Linearizability: A correctness condition for concurrent objects. *ACM Trans. Program. Lang. Syst.*, 12(3):463–492, 1990.

[11] C. A. R. Hoare. *Communicating Sequential Processes*. Prentice Hall, Englewood Cliffs, NJ, 1985.

[12] C. M. Kirsch, H. Payer, H. Röck, and A. Sokolova. Performance, scalability, and semantics of concurrent fifo queues. In *International Conference on Algorithms and Architectures for Parallel Processing*, pages 273–287, 2012.

[13] L. Lamport. How to make a multiprocessor computer that correctly executes multiprocess programs. *IEEE Trans. Computers*, 28(9):690–691, 1979.

[14] Y. Liu, W. Chen, Y. Liu, S. Zhang, J. Sun, and J. S. Dong. Verifying linearizability via optimized refinement checking. *IEEE Transactions on Software Engineering*, 2013.

[15] Y. Liu, W. Chen, Y. A. Liu, and J. Sun. Model checking linearizability via refinement. In *Proceedings of the 2nd World Congress on Formal Methods*, pages 321–337, Berlin, Heidelberg, 2009. Springer-Verlag.

[16] N. Lynch. *Distributed Algorithms*. Morgan Kaufmann, 1997.

[17] H. Payer, H. Röck, C. M. Kirsch, and A. Sokolova. Scalability versus semantics of concurrent fifo queues. In *ACM Symposium on Principles of Distributed Computing*, pages 331–332, 2011.

[18] C. Sadowski, S. N. Freund, and C. Flanagan. Singletrack: A dynamic determinism checker for multithreaded programs. In *European Symposium on Programming*, pages 394–409, 2009.

[19] J. Sun, Y. Liu, J. S. Dong, and C. Chen. Integrating specification and programs for system modeling and verification. In *Proceedings of the third IEEE International Symposium on Theoretical Aspects of Software Engineering*, pages 127–135, 2009.

[20] J. Sun, Y. Liu, J. S. Dong, and J. Pang. PAT: Towards Flexible Verification under Fairness. In *Proceedings of the 21th International Conference on Computer Aided Verification*, volume 5643 of *Lecture Notes in Computer Science*, pages 709–714, 2009.

[21] R. K. Treiber. Systems Programming: Coping with Parallelism. Technical Report RJ 5118, IBM Almaden Research Center, 1986.

[22] V. Vafeiadis. Shape-value abstraction for verifying linearizability. In *International Conference on Verification, Model Checking, and Abstract Interpretation*, pages 335–348, Berlin, Heidelberg, 2009. Springer-Verlag.

[23] V. Vafeiadis, M. Herlihy, T. Hoare, and M. Shapiro. Proving correctness of highly-concurrent linearisable objects. In *ACM SIGPLAN symposium on Principles and practice of parallel programming*, pages 129–136, New York, NY, USA, 2006. ACM.

[24] M. T. Vechev, E. Yahav, and G. Yorsh. Experience with model checking linearizability. In *International SPIN Workshop on Model Checking Software*, pages 261–278, 2009.

[25] W. Vogels. Eventually consistent. *Commun. ACM*, 52(1):40–44, 2009.

[26] C. Wang, R. Limaye, M. Ganai, and A. Gupta. Trace-based symbolic analysis for atomicity violations. In *International Conference on Tools and Algorithms for Construction and Analysis of Systems*, pages 328–342, 2010.

[27] L. Wang and S. D. Stoller. Runtime analysis of atomicity for multithreaded programs. *IEEE Trans. Software Eng.*, 32(2):93–110, 2006.