Simpler Specifications and Easier Proofs of Distributed Algorithms Using History Variables

Saksham Chand and Yanhong A. Liu

Stony Brook University, Stony Brook, NY 11794, USA
{schand, liu}@cs.stonybrook.edu

December 20, 2018

Abstract

This paper studies specifications and proofs of distributed algorithms when only message history variables are used, using Basic Paxos and Multi-Paxos for distributed consensus as precise case studies. We show that not using and maintaining other state variables yields simpler specifications that are more declarative and easier to understand. It also allows easier proofs to be developed by needing fewer invariants and facilitating proof derivations. Furthermore, the proofs are mechanically checked more efficiently.

We show that specifications in TLA$^+$ and proofs in TLA$^+$ Proof System (TLAPS) are reduced by 25% and 27%, respectively, for Basic Paxos, and 46% (from about 100 lines to about 50 lines) and 48% (from about 1000 lines to about 500 lines), respectively, for Multi-Paxos. Overall we need 54% fewer manually written invariants and our proofs have 46% fewer obligations. Our proof for Basic Paxos takes 26% less time for TLAPS to check, and our proofs for Multi-Paxos are checked within 1.5 minutes whereas prior proofs fail to be checked by TLAPS.

Keywords. Distributed algorithms, Specifications, Formal verification, Invariants

1 Introduction

Reasoning about correctness of distributed algorithms is notoriously difficult due to a number of reasons including concurrency, asynchronous networks, unbounded delay, and arbitrary failures. Emerging technologies like autonomous cars are bringing vehicular clouds closer to reality [10], decentralized digital currencies are gathering more attention from academia and industry than ever [34], and with the explosion in the number of nano- and pico- satellites being launched, a similar trend is expected in the field of space exploration as well [32]. All of these systems deal with critical resources like human life, currency, and intricate machinery. This only amplifies the need for employing formal methods to guarantee their correctness.

Verification of distributed algorithms continues to pose a demanding challenge to computer scientists, exacerbated by the fact that paper proofs of these algorithms cannot be trusted [36]. The usual line of reasoning in static analysis of such systems involves manually writing invariants and then using theorem provers to verify that the invariants follow from the specification and that they imply correctness.

History variables and derived variables. A distributed system comprises a set of processes communicating with each other by message passing while performing local actions that may be triggered upon receiving a set of messages and may conclude with sending a set of messages [15, 17]. As such, data processed by any distributed process fall into two categories: (i) *History Variables:* Sets of all messages sent and received[1] and (ii) *Derived Variables:* Local data maintained for efficient computation. Derived variables are often used to maintain results of aggregate queries over sent and received messages.

While reading and writing pseudocode, derived variables are helpful because instead of writing the definition of the variable everywhere, the variable is used instead. Human readers would recall the definition and convince themselves how the algorithm works. While this approach works well for humans, the same is not true for provers. For specifications written with derived variables, invariants have to be added to their proofs which, at the very least, establish that the derived variable implements its definition.

One reason to use derived variables in formal specifications is their existence in pseudocode. Another reason is the lack of high-level languages that provide elegant support for quantifications, history variables, and automatic optimal maintenance of aggregate queries over history variables. The barrier of lack of executable language support for such richness is overcome by high-level languages like DistAlgo [25], which provides native support for history variables, quantifications, and aggregate queries. This motivated us to dispense with derived variables, and study specifications written with only history variables and the impact of this change on their proofs.

Note that uses of history variables provide higher-level specifications of systems in terms of what to compute, as opposed to how to compute with employing and updating derived variables. It makes proofs easier, independent of the logics used for doing the proofs, because important invariants are captured directly in the specifications, rather than hidden under all the incremental updates. On the other hand, it can make model checking much less efficient, just as it can make straightforward execution much less efficient. This is not only because high-level queries are time consuming, but also because maintaining history variables can blow up the state space. This is why automatic incrementalization [30, 31, 11, 24] is essential for efficient implementations, including implementations of distributed algorithms [26, 23]. The same transformations for incrementalization can drastically speed up both program execution and model checking.

This paper. We first describe a systematic style to write specifications of distributed algorithms using message history variables. The only variables in these specifications are the sets of sent and received messages. We show (i) how these are different from the usual pseudocode, (ii) why these are sufficient for specifying all distributed algorithms, and (iii) when these are better for the provers than other specifications. A method is then explained which, given such specifications, allows us to systematically derive many important invariants which are needed to prove correctness. This method exploits the fact that the sets of sent and received messages grow monotonically — messages can only be added or read from these sets, not updated or deleted.

We use three existing specifications and their Safety proofs as our case studies: (i) Basic Paxos for single-valued consensus by Lamport et al., distributed as an example with the TLA$^+$ Proof System (TLAPS) [22], (ii) Multi-Paxos for multi-valued consensus [3], and

[1] This is different from some other references of the term history variables which include sequences of local actions, i.e., execution history [7]

(iii) Multi-Paxos with preemption [3]. Paxos is chosen because it is famous for being a difficult algorithm to grasp, while at the same time it is the core algorithm for distributed consensus—the most fundamental problem in distributed computing. We show that our approach led to significantly reduced sizes of specifications and proofs, numbers of needed manually written invariants, and proof checking times. Our specifications and proofs are available at https://github.com/sachand/HistVar.

This paper is an extended version of [2]. The main extensions are added proof descriptions in Section 3 and complete, cleaned up specification, invariants, and proof in Appendices A, B, and C, respectively.

The rest of the paper is organized as follows. Section 2 details our style of writing specifications using Basic Paxos as an example. We then describe our strategy to systematically derive invariants in Section 3 while also showing how using history variables leads to needing fewer invariants. We discuss Multi-Paxos briefly in Section 4. Results comparing our specifications and proofs with those that do not use history variables are detailed in Section 5. Section 6 discusses related work and concludes.

2 Specifications using message history variables

We demonstrate our approach by developing a specification of Basic Paxos in which we only maintain the set of sent messages. This specification is made to correspond to the specification of Basic Paxos in TLA$^+$ written by Lamport et al. [22]. This is done intentionally to better understand the applicability of our approach. We also simultaneously show Lamport's description of the algorithm in English [19] to aid the comparison, except we rename message types and variable names to match those in his TLA$^+$ specification: *prepare* and *accept* messages are renamed 1a and 2a respectively, their responses are renamed 1b and 2b, respectively, and variable n is renamed b and *bal* in different places.

Distributed consensus. The basic consensus problem, called single-value consensus or single-decree consensus, is to ensure that at most a single value is chosen from among the values proposed by the processes. It is formally defined as

$$Safe \triangleq \forall\, v1, v2 \in \mathcal{V} : Chosen(v1) \land Chosen(v2) \Rightarrow v1 = v2 \tag{1}$$

where $\mathcal{V}$ is the set of possible proposed values, and *Chosen* is a predicate that given a value v evaluates to true iff v was chosen by the algorithm. The specification of *Chosen* is part of the algorithm.

Basic Paxos. Paxos solves the problem of consensus. Two main roles of the algorithm are performed by two kinds of processes:

- $\mathcal{P}$, the set of proposers that propose values that can be chosen.
- $\mathcal{A}$, the set of acceptors that vote for proposed values. A value is chosen when there are enough votes for it.

A set $\mathcal{Q}$ of subsets of the acceptors, that is $\mathcal{Q} \subseteq 2^{\mathcal{A}}$, is used as a quorum system. It must satisfy the following properties:

- $\mathcal{Q}$ is a set cover for $\mathcal{A}$ — $\bigcup_{Q \in \mathcal{Q}} Q = \mathcal{A}$.

- Any two quorums overlap — $\forall Q1, Q2 \in \mathcal{Q} : Q1 \cap Q2 \neq \emptyset$.

The most commonly used quorum system takes any majority of acceptors as an element in $\mathcal{Q}$. For example, if $\mathcal{A} = \{1, 2, 3\}$, then the majority based quorum set is $\mathcal{Q} = \{\{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}\}$. Quorums are needed because the system can have arbitrary failures. If a process waits for replies from all other processes, as in Two-Phase Commit, the system will hang in the presence of even one failed process. For instance, in the mentioned example, the system will continue to work even if acceptor 3 fails because at least one quorum, which is $\{1, 2\}$, is alive.

Basic Paxos solves the problem of single-value consensus. It defines predicate *Chosen* as

$$Chosen(v) \triangleq \exists\, Q \in \mathcal{Q} : \forall\, a \in Q : \exists\, b \in \mathcal{B} : sent(\text{“2b”}, a, b, v) \tag{2}$$

where $\mathcal{B}$ is the set of proposal numbers, also known as ballot numbers [18], which is any set that can be strictly totally ordered. $sent(\text{“2b”}, a, b, v)$ means that a message of type 2b with ballot number b and value v was sent by acceptor a (to some set of processes). An acceptor votes by sending such a message.

TLA$^+$. The specifications presented in this paper are written in TLA$^+$, an extension of the Temporal Logic of Actions (TLA) [17], a logic for specifying concurrent and distributed programs and reasoning about their properties. In TLA, a *state* is an instantiation of the variables of the program to values. An *action* is a relation between a current state and a new state, specifying the effect of executing a sequence of instructions. For example, the instruction $x := x + 1$ is represented in TLA and TLA$^+$ by the action $x' = x + 1$. An action is represented by a formula over unprimed and primed variables where unprimed variables refer to the values of the variables in the current state and primed variables refer to the values of the variables in the new state.

A program is specified by its actions and initial states. Formally, a program is specified as $Spec \triangleq Init \wedge \Box[Next]_{vars}$ where *Init* is a predicate that holds for initial states of the program, *Next* is a disjunction of all the actions of the program, and *vars* is the tuple of all the variables. The expression $[Next]_{vars}$ is true if either *Next* is true, implying some action is true and therefore executed, or *vars* stutters, that is, the values of the variables are same in the current and new states. $\Box$ is the temporal operator *always*.

As a simple example, consider this specification of a clock based on Lamport's logical clock [16] but on a shared memory system:

$$\begin{aligned}
&\text{VARIABLE } c \\
&Max(S) \triangleq \text{CHOOSE } e \in S : \forall\, f \in S : e \geq f \\
&Init \triangleq c = [p \in \{0, 1\} \mapsto 0] \\
&LocalEvent(p) \triangleq c' = [c \text{ EXCEPT } ![p] = c[p] + 1] \\
&ReceiveEvent(p) \triangleq c' = [c \text{ EXCEPT } ![p] = Max(\{c[p], c[1 - p]\}) + 1] \\
&Next \triangleq \exists\, p \in \{0, 1\} : LocalEvent(p) \vee ReceiveEvent(p) \\
&Spec \triangleq Init \wedge \Box[Next]_{\langle c \rangle}
\end{aligned} \tag{3}$$

The system has two processes numbered 0 and 1. Variable c stores their current clock values as a function from process numbers to clock values. Both processes start with clock value 0, as specified in *Init*. $LocalEvent(p)$ specifies that process p has executed some local action and therefore increments its clock value. The expression $c' = [c \text{ EXCEPT } ![p] = c[p]+1]$ means that function c' is the same as function c except that $c'[p]$ is $c[p]+1$. $ReceiveEvent(p)$

Phase 1a. A proposer selects a proposal number b and sends a 1a request with number b to a majority of acceptors.	
Lamport et al.'s	Using *sent* only
$Phase1a(b \in \mathcal{B}) \triangleq$ $\wedge \nexists\, m \in sent : (m.type = \text{"1a"}) \wedge (m.bal = b)$ $\wedge Send([type \mapsto \text{"1a"}, bal \mapsto b])$ $\wedge$ UNCHANGED $\langle maxVBal, maxBal, maxVal \rangle$	$Phase1a(b \in \mathcal{B}) \triangleq$ $Send([type \mapsto \text{"1a"}, bal \mapsto b])$

Figure 1: Phase 1a of Basic Paxos

specifies that process p updates its clock value to 1 greater than the higher of its and the other process' clock value. We define operator *Max* to obtain the highest of a set of values. CHOOSE denotes Hilbert's ϵ operator that returns some nondeterministically chosen term satisfying the body of the CHOOSE expression if it exists, otherwise an error is raised.

Basic Paxos variables. Lamport et al.'s specification of Basic Paxos has four global variables.

- *msgs*— history variable maintaining the set of messages that have been sent. Processes read from or add to this set but cannot remove from it. We rename this to *sent* in both ours and Lamport et al.'s specifications for clarity purposes. This is the only variable maintained in our specifications.
- *maxBal*—for each acceptor, the highest ballot seen by it.
- *maxVBal* and *maxVal*—for each acceptor, *maxVBal* is the highest ballot in which it has voted, and *maxVal* is the value it voted for in that ballot.

Basic Paxos algorithm steps. The algorithm consists of repeatedly executing two phases. Each phase comprises two actions, one by acceptors and one by proposers.

- **Phase 1a.** Fig. 1 shows Lamport's description in English followed by Lamport et al.'s and our specifications in TLA$^+$. *Send* is an operator that adds its argument to *sent*, i.e., $Send(m) \triangleq sent' = sent \cup \{m\}$.
  1. The first conjunct in Lamport et al.'s specification is not mentioned in the English description and is not needed. Therefore it was removed.
  2. The third conjunct is also removed because the only variable our specification maintains is *sent*, which is updated by *Send*.
- **Phase 1b.** Fig. 2 shows the English description and the specifications of Phase 1b. The first two conjuncts in both specifications capture the precondition in the English description. The remaining conjuncts specify the action.
  1. The first conjunct states that message m received by acceptor a is of type 1a.
  2. The second conjunct ensures that the proposal number *bal* in the 1a message m is higher than that of any 1a request responded to by a. In Lamport et al.'s specification, derived variable $maxBal[a]$ maintains the highest proposal number that a has ever responded to, in 1b and 2b messages, and its second conjunct

Phase 1b. If an acceptor receives a 1a request with number *bal* greater than that of any 1a request to which it has already responded, then it responds to the request with a promise not to accept any more proposals numbered less than *bal* and with the highest-numbered proposal (if any) that it has accepted.

Lamport et al.'s	Using *sent* only
$Phase1b(a \in \mathcal{A}) \triangleq$ $\exists\, m \in sent :$ $\land m.type = \text{"1a"}$ $\land m.bal > maxBal[a]$ $\land Send([type \mapsto \text{"1b"},$ $acc \mapsto a, bal \mapsto m.bal,$ $maxVBal \mapsto maxVBal[a],$ $maxVal \mapsto maxVal[a]])$ $\land maxBal' =$ $[maxBal \text{ EXCEPT } ![a] = m.bal]$ $\land \text{UNCHANGED } \langle maxVBal, maxVal \rangle$	$Phase1b(a \in \mathcal{A}) \triangleq$ $\exists\, m \in sent, r \in max\_prop(a) :$ $\land m.type = \text{"1a"}$ $\land \forall\, m2 \in sent : m2.type \in \{\text{"1b"}, \text{"2b"}\} \land$ $m2.acc = a \Rightarrow m.bal > m2.bal$ $\land Send([type \mapsto \text{"1b"},$ $acc \mapsto a, bal \mapsto m.bal,$ $maxVBal \mapsto r.bal,$ $maxVal \mapsto r.val)$ $2bs(a) \triangleq \{m \in sent : m.type = \text{"2b"} \land m.acc = a\}$ $max\_prop(a) \triangleq$ IF $2bs(a) = \emptyset$ THEN $\{[bal \mapsto -1, val \mapsto \bot]\}$ ELSE $\{m \in 2bs(a) : \forall m2 \in 2bs(a) : m.bal \geq m2.bal\}$

Figure 2: Phase 1b of Basic Paxos

uses $m.bal > maxBal[a]$. Using *sent* only, we capture this intent more directly, as $\forall\, m2 \in sent : m2.type \in \{\text{"1b"}, \text{"2b"}\} \land m2.acc = a \Rightarrow m.bal > m2.bal$, because those $m2$'s are the response messages that a has ever sent.

3. The third conjunct is the action of sending a promise (1b message) not to accept any more proposals numbered less than *bal* and with the highest-numbered proposal (if any) that a has accepted, i.e., has sent a 2b message. This proposal is maintained in Lamport et al.'s specification in derived variables *maxVBal* and *maxVal*. We specify this proposal as $max\_prop(a)$, which is either the set of proposals that have the highest proposal number among all accepted by a or if a has not accepted anything, then $\{[bal \mapsto -1, val \mapsto \bot]\}$ as the default, where $-1 \notin \mathcal{B}$ and is smaller than all ballots and $\bot \notin \mathcal{V}$. This corresponds to initialization in Lamport et al.'s specification as shown in Fig. 5. Note that the specification in Appendix A writes $\bot$ as *None*.

4. The remaining conjuncts in Lamport et al.'s specification maintain the variable $maxBal[a]$. A compiler that implements incrementalization [24] over queries would automatically generate and maintain such a derived variable to optimize the corresponding query.

- **Phase 2a.** Fig. 3 shows Phase 2a. The specifications differ from the English description by using a set of quorums, $\mathcal{Q}$, instead of a majority. The only difference between the two specifications is the removed UNCHANGED conjunct when using *sent* only. It is important to note that the English description fails to mention the first conjunct—a conjunct without which the specification is unsafe. Every 2a message must have a unique ballot.

Note that the first conjunct in Lamport et al.'s specification (and therefore ours as well) states that none of the 2a messages sent so far has *bal* equal to b. This is not directly implementable in a real system because this quantification query requires accessing message histories of all processes. We leave this query as is for two main reasons: (i) The focus of this paper is to demonstrate the use of history variables against derived variables and compare them in the light of simpler specification and verification. This removes derived variables but leaves queries on history variables unchanged even though they are not directly implementable. (ii) There is a commonly-used, straightforward, efficient way to implement this query - namely realizing ballot as a tuple in $\mathbb{N} \times \mathcal{P}$ [18]. So a proposer only executes Phase 2a on a ballot proposed by itself (i.e., sent a 1a message with that ballot) and, for efficient implementation, only executes Phase 2a on the highest ballot that it has proposed.

Phase 2a. If the proposer receives a response to its 1a requests (numbered b) from a majority of acceptors, then it sends a 2a request to each of those acceptors for a proposal numbered b with a value v, where v is the value of the highest-numbered proposal among the 1b responses, or is any value if the responses reported no proposals.

Lamport et al.'s

$$
\begin{array}{l}
Phase2a(b \in \mathcal{B}) \triangleq \\
\land \nexists m \in sent : m.type = \text{"2a"} \land m.bal = b \\
\land \exists\, v \in \mathcal{V}, Q \in \mathcal{Q}, S \subseteq \{m \in sent : \\
\quad m.type = \text{"1b"} \land m.bal = b\} : \\
\qquad \land \forall\, a \in Q : \exists\, m \in S : m.acc = a \\
\qquad \land \lor \forall\, m \in S : m.maxVBal = -1 \\
\qquad\quad \lor \exists\, c \in 0..(b-1) : \\
\qquad\qquad \land \forall\, m \in S : m.maxVBal \leq c \\
\qquad\qquad \land \exists\, m \in S : \land m.maxVBal = c \\
\qquad\qquad\qquad \land m.maxVal = v \\
\quad \land Send([type \mapsto \text{"2a"}, bal \mapsto b, val \mapsto v]) \\
\land \text{UNCHANGED } \langle maxBal, maxVBal, \\
\quad maxVal \rangle
\end{array}
$$

Using *sent* only

$$
\begin{array}{l}
Phase2a(b \in \mathcal{B}) \triangleq \\
\land \nexists m \in sent : m.type = \text{"2a"} \land m.bal = b \\
\land \exists\, v \in \mathcal{V}, Q \in \mathcal{Q}, S \subseteq \{m \in sent : \\
\quad m.type = \text{"1b"} \land m.bal = b\} : \\
\qquad \land \forall\, a \in Q : \exists\, m \in S : m.acc = a \\
\qquad \land \lor \forall\, m \in S : m.maxVBal = -1 \\
\qquad\quad \lor \exists\, c \in 0..(b-1) : \\
\qquad\qquad \land \forall\, m \in S : m.maxVBal \leq c \\
\qquad\qquad \land \exists\, m \in S : \land m.maxVBal = c \\
\qquad\qquad\qquad \land m.maxVal = v \\
\quad \land Send([type \mapsto \text{"2a"}, bal \mapsto b, val \mapsto v])
\end{array}
$$

Figure 3: Phase 2a of Basic Paxos

- **Phase 2b.** Fig. 4 shows Phase 2b. Like Phase 1b, we replace the second conjunct with the corresponding query over *sent* and remove updates to the derived variables.

Overall Basic Paxos algorithm. To complete the algorithm specification, we define, and compare, *vars*, *Init*, *Next*, and *Spec* which are typical TLA$^+$ operator names for the set of variables, the initial state, possible actions leading to the next state, and the system specification, respectively, in Fig. 5.

Lamport et al.'s initialization of *maxVBal* and *maxVal* to -1 and $\perp$, respectively, is moved to our definition of *max_prop* in Fig. 2. We do not need initialization of *maxBal* because if no 1b or 2b messages have been sent, the universally quantified queries over them would be vacuously true. In Lamport et al.'s specification, this is achieved by initializing *maxBal* to -1, which is smaller than all ballots, and thus, the conjunct $m.bal > maxBal[a]$ in Fig. 2 holds for the first 1a message received.

Phase 2b. If an acceptor receives a 2a request for a proposal numbered *bal*, it accepts the proposal unless it has already responded to a 1a request having a number greater than *bal*.

Lamport et al.'s	Using *sent* only
$Phase2b(a \in \mathcal{A}) \triangleq$ $\exists\, m \in sent :$ $\wedge m.type = \text{“2a”}$ $\wedge m.bal \geq maxBal[a]$ $\wedge Send([type \mapsto \text{“2b”}, acc \mapsto a,$ $\quad bal \mapsto m.bal, val \mapsto m.val])$ $\wedge maxBal' = [maxBal \text{ EXCEPT } ![a] = m.bal]$ $\wedge maxVBal' = [maxVBal \text{ EXCEPT } ![a] = m.bal]$ $\wedge maxVal' = [maxVal \text{ EXCEPT } ![a] = m.val]$	$Phase2b(a \in \mathcal{A}) \triangleq$ $\exists\, m \in sent :$ $\wedge m.type = \text{“2a”}$ $\wedge \forall\, m2 \in sent : m2.type \in \{\text{“1b”}, \text{“2b”}\} \wedge$ $\quad m2.acc = a \Rightarrow m.bal \geq m2.bal$ $\wedge Send([type \mapsto \text{“2b”}, acc \mapsto a,$ $\quad bal \mapsto m.bal, val \mapsto m.val])$

Figure 4: Phase 2b of Basic Paxos

The complete Basic Paxos algorithm specification is given in Appendix A.

Lamport et al.'s	Using *sent* only
$vars \triangleq \langle sent, maxBal, maxVBal, maxVal \rangle$ $Init \triangleq \wedge sent = \emptyset$ $\wedge maxVBal = [a \in \mathcal{A} \mapsto -1]$ $\wedge maxBal = [a \in \mathcal{A} \mapsto -1]$ $\wedge maxVal = [a \in \mathcal{A} \mapsto \bot]$	$vars \triangleq \langle sent \rangle$ $Init \triangleq sent = \emptyset$
$Next \triangleq \vee \exists\, b \in \mathcal{B} : Phase1a(b) \vee Phase2a(b)$ $\vee \exists\, a \in \mathcal{A} : Phase1b(a) \vee Phase2b(a)$ $Spec \triangleq Init \wedge \Box[Next]_{vars}$	

Figure 5: Overall algorithm specification

3 Invariants and proofs using message history variables

Invariants of a distributed algorithm can be categorized into the following three kinds:

1. *Type invariants.* These ensure that all data processed in the algorithm is of valid type. For example, messages of type 1a must have a field $bal \in \mathcal{B}$. If an action sends a 1a message with *bal* missing or $bal \notin \mathcal{B}$, a type invariant is violated.

2. *Message invariants.* These are invariants defined on message history variables. For example, each message of type 2a has a unique *bal*. This is expressed by the invariant $\forall\, m1, m2 \in sent : m1.type = \text{“2a”} \wedge m2.type = \text{“2a”} \wedge m1.bal = m2.bal \Rightarrow m1 = m2$.

3. *Process invariants.* These state properties about the data maintained in derived variables. For example, in Lamport et al.'s specification, one such invariant is that for any acceptor a, $maxBal[a] \geq maxVBal[a]$.

Fig. 6 shows and compares all invariants used in Lamport et al.'s proof vs. ours. The following operators are used in the invariants for brevity:

$$
\begin{array}{l}
\mathit{VotedForIn}(a, v, b) \triangleq \exists\, m \in \mathit{sent} : \\
\quad \land\ m.\mathit{type} = \text{“2b”} \\
\quad \land\ m.\mathit{acc} = a \\
\quad \land\ m.\mathit{val} = v \\
\quad \land\ m.\mathit{bal} = b \\
\mathit{WontVoteIn}(a, b) \triangleq \forall\, v \in \mathcal{V} : \qquad \backslash * \text{ Lamport et al.'s} \\
\quad \land\ \neg\, \mathit{VotedForIn}(a, v, b) \\
\quad \land\ \mathit{maxBal}[a] > b \\
\mathit{WontVoteIn}(a, b) \triangleq \forall\, v \in \mathcal{V} : \qquad \backslash * \text{ Using } \mathit{sent} \text{ only} \\
\quad \land\ \neg\, \mathit{VotedForIn}(a, v, b) \\
\quad \land\ \exists\, m \in \mathit{sent} : \\
\qquad \land\ m.\mathit{type} \in \{\text{“1b”}, \text{“2b”}\} \\
\qquad \land\ m.\mathit{acc} = a \\
\qquad \land\ m.\mathit{bal} > b \\
\mathit{SafeAt}(v, b) \triangleq \forall\, b2 \in 0..(b-1) : \exists\, Q \in \mathcal{Q} : \forall\, a \in Q : \\
\quad \lor\ \mathit{VotedForIn}(a, v, b2) \\
\quad \lor\ \mathit{WontVoteIn}(a, b2)
\end{array}
\tag{4}
$$

The complete invariants, auxiliary operators, and the safety property to be proved can be found in Appendix B.

Type invariants reduced to one. Lamport et al. define four type invariants, one for each variable they maintain. *Messages* is the set of all possible valid messages. We require only one, (I1). This invariant asserts that the type of all sent messages is valid. (I2-4) are not applicable to our specification.

Process invariants not needed. Lamport et al. define four process invariants, (I5-8), regarding variables maxVal, $\mathit{maxVBal}$, and maxBal. They are not applicable to our specification, and need not be given in our proof.

(I5) Because $\mathit{maxBal}[a]$ is the highest ballot ever seen by a and $\mathit{maxVBal}[a]$ is the highest ballot a has voted for, we have the following invariants:

$$
\begin{array}{rl}
\mathit{maxBal}[a] \equiv \mathtt{max}(\{m.\mathit{bal} : m \in \mathit{sent} \land m.\mathit{type} \in \{\text{“1b”},\text{“2b”}\} & \land m.\mathit{acc} = a\}) \\
\mathit{maxVBal}[a] \equiv \mathtt{max}(\{m.\mathit{bal} : m \in \mathit{sent} \land m.\mathit{type} \in \{\text{“2b”}\} & \land m.\mathit{acc} = a\})
\end{array}
\tag{5}
$$

where $\mathtt{max}(S) \triangleq$ CHOOSE $e \in S \cup \{-1\} : \forall\, f \in S : e \geq f$. Note that `max` is not in TLA$^+$ and has to be user-defined. Invariant (I5) is needed in Lamport et al.'s proof but not ours because they use derived variables whereas we specify the properties directly. For example, for Lamport et al.'s Phase 1b, one cannot deduce $m.\mathit{bal} > \mathit{maxVBal}[a]$ without (I5), whereas in our Phase 1b, definitions of $2bs$ and $\mathit{max\_prop}$ along with the second conjunct are enough to deduce it.

	Lamport et al.'s proof	Our proof
Type Invariants	(I1) $sent \subseteq Messages$ (I2) $maxVBal \in [\mathcal{A} \to \mathcal{B} \cup \{-1\}]$ (I3) $maxBal \in [\mathcal{A} \to \mathcal{B} \cup \{-1\}]$ (I4) $maxVal \in [\mathcal{A} \to \mathcal{V} \cup \{\bot\}]$	$sent \subseteq Messages$
Process Invariants $\forall\, a \in \mathcal{A}$	(I5) $maxBal[a] \geq maxVBal[a]$ (I6) $maxVal[a] = \bot \Leftrightarrow maxVBal[a] = -1$ (I7) $maxVBal[a] \geq 0 \Rightarrow$ $VotedForIn(a, maxVal[a], maxVBal[a])$ (I8) $\forall\, b \in \mathcal{B} : b > maxVBal[a] \Rightarrow$ $\nexists v \in \mathcal{V} : VotedForIn(a, v, b)$	
Message Invariants $\forall m \in sent$	(I9) $m.type = \text{"2b"} \Rightarrow m.bal \leq maxVBal[m.acc]$ (I10) $m.type = \text{"1b"} \Rightarrow m.bal \leq maxBal[m.acc]$	
	(I11) $m.type = \text{"1b"} \Rightarrow$ $\lor \land m.maxVal \in \mathcal{V} \land m.maxVBal \in \mathcal{B}$ $\land VotedForIn(m.acc,$ $m.maxVal, m.maxVBal)$ $\lor m.maxVBal = -1 \land m.maxVal = \bot$	$m.type = \text{"1b"} \Rightarrow$ $\lor VotedForIn(m.acc,$ $m.maxVal, m.maxVBal)$ $\lor m.maxVBal = -1$
	(I12) $m.type = \text{"1b"} \Rightarrow$ $\forall\, b2 \in m.maxVBal + 1..m.bal - 1 : \nexists v \in \mathcal{V} : VotedForIn(m.acc, v, b2)$ (I13) $m.type = \text{"2a"} \Rightarrow SafeAt(m.val, m.bal)$ (I14) $m.type = \text{"2a"} \Rightarrow$ $\forall\, m2 \in sent : m2.type = \text{"2a"} \land m2.bal = m.bal \Rightarrow m2 = m$ (I15) $m.type = \text{"2b"} \Rightarrow$ $\exists\, m2 \in sent : m2.type = \text{"2a"} \land m2.bal = m.bal \land m2.val = m.val$	

Figure 6: Comparison of invariants. Our proof does not need I2-I10, and needs only I1, a simpler I11, and I12-I15.

(I6) Lamport et al.'s proof needs this invariant to prove (I11). Because the initial values are part of *Init* and are not explicitly present in their Phase 1b, this additional invariant is needed to carry this information along. We include the initial values when specifying the action in Phase 1b and therefore do not need this invariant.

(I7) This invariant is obvious from the definition of *VotedForIn* in Equation (3) and property of $maxVBal$ in Equation (4). The premise $maxVBal[a] \geq 0$ is needed by Lamport et al.'s proof to differentiate from the initial value -1 of $maxVBal[a]$.

(I8) This states that a has not voted for any value at a ballot greater than $maxVBal[a]$. This invariant need not be manually given in our proofs because it is implied from the definition of $maxVBal[a]$.

Message invariants not needed or more easily proved. Before detailing the message invariants, we present a systematic method that can derive several useful invariants used by Lamport et al. and thus make the proofs easier. This method is based on the following properties of our specifications and distributed algorithms:

1. *sent* grows monotonically, that is, the only operations on it are read and add.

2. Message invariants hold for each sent message of some type, i.e., they are of the form $\forall\, m \in sent : m.type = \tau \Rightarrow \Phi(m)$, or more conveniently if we define $sent_\tau = \{m \in$

$sent : m.type = \tau\}$, we have $\forall\, m \in sent_\tau : \Phi(m)$.

3. $sent = \emptyset$ initially, so the message invariants are vacuously true in the initial state of the system.

4. Distributed algorithms usually implement a logical clock for ordering two arbitrary messages. In Paxos, this is done by ballots.

We demonstrate our method by deriving (I15). The method is applied for each message type used in the algorithm. Invariant (I15) is about 2b messages. We first identify all actions that send 2b messages and then do the following:

1. **Increment.** 2b messages are sent in Phase 2b as specified in Fig. 4. We first determine the increment to $sent$, $\Delta(sent)$, the new messages sent in Phase 2b. We denote a message in $\Delta(sent)$ by δ for brevity. We have, from Fig. 4,

$$\delta = [type \mapsto \text{“2b”}, acc \mapsto a, bal \mapsto m.bal, val \mapsto m.val] \tag{6}$$

2. **Analyze.** We deduce properties about the messages in $\Delta(sent)$. For 2b messages, we deduce the most straightforward property that connects the contents of messages in $\Delta(sent)$ with the message m, from Fig. 4,

$$\phi(\delta) = \exists\, m \in sent : m.type = \text{“2a”} \wedge \delta.bal = m.bal \wedge \delta.val = m.val \tag{7}$$

3. **Integrate.** Because (i) $sent$ monotonically increases, and (ii) ϕ is an existential quantification over $sent$, ϕ holds for all increments to $sent_{2b}$. Property (i) means that once the existential quantification in ϕ holds, it holds forever. Integrating both sides of Equation (6) in the space of 2b messages yields (I15), i.e.,

$$\begin{aligned}\Phi(sent_{2b}) = \forall\, m2 \in sent_{2b} : \exists\, m \in sent :& \\ m.type = \text{“2a”} \wedge m2.bal = m.bal \wedge m2.val = m.val&\end{aligned} \tag{8}$$

The case for ϕ being universally quantified over $sent$ is discussed with invariant (I12).

Other message invariants. (I9) and (I10) follow directly from Equation (4) and need not be manually specified for our proof. We also derive (I11), (I12), and (I14) as described in the following.

(I11) Like (I15), (I11) can also be systematically derived, from our Phase 1b in Fig. 2. This invariant is less obvious when variables $maxVal$ and $maxVBal$ are explicitly used and updated because (i) they are not updated in the same action that uses them, requiring additional invariants to carry their meaning to the proofs involving the actions that use them, and (ii) it is not immediately clear if these variables are being updated in Lamport et al.'s Phase 2b in Fig. 4 because a 2b message is being sent or because a 2a message was received.

(I12) To derive (I11) and (I15), we focused on *where* the contents of the new message come from. For (I12), we analyze *why* those contents were chosen. From our Phase 1b with definitions of $2bs$ and max_prop in Fig. 2, we have

$$\begin{aligned}
&\phi(\delta) = \\
&\quad \vee \wedge \exists\, m \in sent :\ m.type = \text{“2b”} \wedge m.acc = \delta.acc \\
&\qquad \wedge \forall\, m \in sent :\ m.type = \text{“2b”} \wedge m.acc = \delta.acc \Rightarrow \\
&\qquad\qquad \delta.maxVBal \geq m.bal \\
&\quad \vee \wedge \nexists m \in sent :\ m.type = \text{“2b”} \wedge m.acc = \delta.acc \\
&\qquad \wedge \delta.maxVBal = -1
\end{aligned} \tag{9}$$

ϕ has two disjuncts—the first has a universal quantification and the second has negated existential, which is universal in disguise. If *sent* is universally quantified, integration like for (I15) is not possible because the quantification only holds *at the time of the action*. As new messages are sent in the future, the universal may become violated.

The key is the phrase *at the time*. One way to work around the universal is to add a time field in each message and update it in every action as a message is sent, like using a logical clock. Then, a property like $\phi(\delta) = \forall\, m \in sent_\tau : \psi(m)$ can be integrated to obtain

$$\Phi(sent_\tau) = \forall\, m2 \in sent_\tau : \forall\, m \in sent : m.time < m2.time \Rightarrow \psi(m) \tag{10}$$

Because ballots act as the logical clock in Paxos, we do not need to specify a separate logical clock and we can perform the above integration on Equation (8) to obtain the invariant (I12).

(I14) This invariant is of the form $\forall\, m1, m2 \in sent_\tau, t : \psi(m1, t) \wedge \psi(m2, t) \Rightarrow m1 = m2$. In this case, $\psi(m, t) \triangleq m.bal = t$. Deriving invariants like (I14) is nontrivial unless ψ is already known. In some cases, ψ can be guessed. The intuition is to look for a universal quantification (or negated existential) in the specification of an action. The ideal case is when the quantification is on the message type being sent in the action. Potential candidates for ψ may be hidden in such quantifications. Moreover, if message history variables are used, these quantifications are easier to identify.

Starting with a guess of ψ, we identify the change in the counting measure (cardinality) of the set $\{t : m \in sent_\tau \wedge \psi(m, t)\}$ along with that of $sent_\tau$. In the case of (I14), we look for $\Delta(|\{m.bal : m \in sent_{2a}\}|)$. From our Phase 2a in Fig. 3, we have

$$\begin{aligned}
&\Delta(\{m.bal : m \in sent_{2a}\}) = \{b\} \\
&\phi(b) = \nexists m \in sent : m.type = \text{“2a”} \wedge m.bal = b, \\
&\quad \text{where } b \in \Delta(\{m.bal : m \in sent_{2a}\})
\end{aligned} \tag{11}$$

Rewriting ϕ as $\{b\} \not\subseteq \{m.bal : m \in sent_{2a}\}$, it becomes clear that $\Delta(|\{m.bal : m \in sent_{2a}\}|) = 1$. Meanwhile, $\Delta(|\{m \in sent_{2a}\}|) = 1$. Because the counting measure increases by the same amount for both, (I14) can be derived safely.

TLAPS. The proofs presented in this paper are written and checked in TLA$^+$ Proof System (TLAPS), a tool that mechanically checks proofs of properties of systems specified in TLA$^+$. Proofs are written in a hierarchical style [20], and are transformed to individual proof obligations that are sent to backend theorem provers. The primary backend provers are Isabelle

and Zenon, with the SMT solvers CVC3, Z3, veriT, and Yices as backups. Temporal formulas are proved using LS4, a PTL (Propsitional Temporal Logic) prover. Users can specify which prover they want to use by using its name and can specify the timeout for each obligation separately.

As an example, we present the proof of a simple type invariant about the clock specification in (3) — It is always the case that $c \in [\{0,1\} \to \mathbb{N}]$:

$$\begin{array}{l} \mathit{TypeOK} \triangleq c \in [\{0,1\} \to \mathbb{N}] \\ \text{THEOREM } \mathit{Inv} \triangleq \mathit{Spec} \Rightarrow \Box \mathit{TypeOK} \\ \langle 1 \rangle. \text{ USE DEF } \mathit{TypeOK} \\ \langle 1 \rangle 1. \mathit{Init} \Rightarrow \mathit{TypeOK} \text{ BY DEF } \mathit{Init} \\ \langle 1 \rangle 2. \mathit{TypeOK} \wedge [\mathit{Next}]_{\langle c \rangle} \Rightarrow \mathit{TypeOK}' \\ \qquad \text{BY DEF } \mathit{Next}, \mathit{LocalEvent}, \mathit{ReceiveEvent} \\ \langle 1 \rangle. \text{ QED BY } \langle 1 \rangle 1, \langle 1 \rangle 2, \texttt{PTL} \text{ DEF } \mathit{Spec} \end{array} \tag{12}$$

The proof of theorem *Inv* is written in a hierarchical fashion. It is proved by two steps, named $\langle 1 \rangle 1$ and $\langle 1 \rangle 2$, and RuleINV1 by Lamport [17]. Proof steps in TLAPS are typically written as:

$$\langle x \rangle y.\ \mathit{Assertion} \text{ BY } e_1, \ldots, e_m \text{ DEF } d_1, \ldots, d_n \tag{13}$$

which states that step number $\langle x \rangle y$ proves *Assertion* by assuming $e_1, \ldots, e_m$, and expanding the definitions of $d_1, \ldots, d_n$. For example, step $\langle 1 \rangle 1$ proves $\mathit{Init} \Rightarrow \mathit{TypeOK}$ by expanding the definition of *Init*. The QED step for $\langle 1 \rangle$ requires us to invoke a PTL prover because *Inv* is a temporal formula.

Basic Paxos Proof. The main property to prove is *Safety*, defined as follows:

$$\begin{array}{l} \mathit{Inv} \triangleq \mathit{TypeOK} \wedge \mathit{MsgInv} \\ \mathit{Safe} \triangleq \forall\, v1, v2 \in \mathcal{V} : \mathit{Chosen}(v1) \wedge \mathit{Chosen}(v2) \Rightarrow v1 = v2 \\ \text{THEOREM } \mathit{Safety} \triangleq \mathit{Spec} \Rightarrow \Box \mathit{Safe} \end{array} \tag{14}$$

To proceed, we first define *Inv* and prove $\mathit{Inv} \Rightarrow \mathit{Safe}$. We then prove $\mathit{Spec} \Rightarrow \Box \mathit{Inv}$ and, by temporal logic, conclude $\mathit{Spec} \Rightarrow \Box \mathit{Safe}$. Note that property *Safety* is called *Consistent*, and invariant *Safe* is called *Consistency* by Lamport et al. [22].

To prove the *Safe* property for the algorithm, we first prove the following helper lemmas for three important properties:

1. Lemma *VotedInv*. If any acceptor votes any pair $\langle v, b \rangle$, then the predicate $\mathit{SafeAt}(v, b)$ holds:

$$\begin{array}{l} \text{LEMMA } \mathit{VotedInv} \triangleq \mathit{MsgInv} \wedge \mathit{TypeOK} \Rightarrow \forall\, a \in \mathcal{A}, v \in \mathcal{V}, b \in \mathcal{B} : \\ \qquad \mathit{VotedForIn}(a, v, b) \Rightarrow \mathit{SafeAt}(v, b) \end{array} \tag{15}$$

2. Lemma *VotedOnce*. If acceptor $a1$ votes pair $\langle v1, b \rangle$ and acceptor $a2$ votes pair $\langle v2, b \rangle$, then $v1 = v2$:

$$\begin{array}{l} \text{LEMMA } \mathit{VotedOnce} \triangleq \mathit{MsgInv} \Rightarrow \forall\, a1, a2 \in \mathcal{A}, v1, v2 \in \mathcal{V}, b \in \mathcal{B} : \\ \qquad \mathit{VotedForIn}(a1, v1, b) \wedge \mathit{VotedForIn}(a2, v2, b) \Rightarrow v1 = v2 \end{array} \tag{16}$$

3. Lemma *SafeAtStable*. If pair $\langle v, b \rangle$ is safe in the current state, it remains safe in the next state, where state transition is defined by *Next*.

$$\begin{aligned} &\text{LEMMA } \mathit{SafeAtStable} \triangleq \mathit{Inv} \wedge \mathit{Next} \Rightarrow \forall\, v \in \mathcal{V}, b \in \mathcal{B} : \\ &\quad \mathit{SafeAt}(v, b) \Rightarrow \mathit{SafeAt}(v, b)' \end{aligned} \tag{17}$$

The proof of $\mathit{Spec} \Rightarrow \Box \mathit{Inv}$ follows the same strategy used in Chand et al. [3]. The proof is inductive, written in a hierarchical style [20]. The base case proves $\mathit{Init} \Rightarrow \mathit{Inv}$. The inductive case considers each action in *Next* individually, and proves that *Inv* holds in the next state given that it holds in the current state.

The complete proof for Basic Paxos spans about 2 pages, and is summarized as follows:

1. Helper lemmas and their proofs are about half a page. This includes proofs for lemmas *VotedInv*, *VotedOnce*, and *SafeAtStable*.

2. The proof of type invariant *TypeOK* is a quarter page, using only a 1-level proof for each action.

3. The proof of message invariant *MsgInv* is less than a page, using 1-level proofs for actions $\mathit{Phase1a}$, $\mathit{Phase1b}$, and $\mathit{Phase2b}$, taking less than a half a page altogether, and a 4-level proof for $\mathit{Phase2a}$, taking half a page.

4. The proof of theorem *Safety* using $\mathit{Spec} \Rightarrow \Box \mathit{Inv}$ is a quarter page, with a straightforward argument of $\mathit{Inv} \Rightarrow \mathit{Safe}$.

The complete TLAPS-checked proof is given in Appendix C.

4 Multi-Paxos

Multi-Paxos specification. We have developed new specifications of Multi-Paxos and Multi-Paxos with Preemption that use only message history variables, by removing derived variables from the specifications described in Chand et al. [3]. This is done in a way similar to how we removed derived variables from Lamport et al.'s specification of Basic Paxos.

The most interesting action here was preemption. With preemption, if an acceptor receives a 1a or 2a message with *bal* smaller than the highest that it has seen, it responds with a preempt message that contains the highest ballot seen by the acceptor. Upon receiving such a message, the receiving proposer would pick a new ballot that is higher than the ballots of all received preempt messages.

This is a good opportunity to introduce the other message history variable, *received*, the set of all messages received. It is different from *sent* because a message could be delayed indefinitely before being received, if at all. In [3], derived variable *proBallot* is introduced to maintain the result of this query on received messages. We contrast this with our new specification in Fig. 7. $\mathit{Receive}(m)$ adds message m to *received*, i.e., $\mathit{Receive}(m) \triangleq \mathit{received}' = \mathit{received} \cup \{m\}$.

Multi-Paxos proof. While we observed a 27% decrease in proof size for Basic Paxos, for Multi-Paxos this decrease was 48%. Apart from the points described in Section 3, an

Chand et al. [3]	Using *sent* and *received*
$NewBallot(c \in \mathcal{B}) \triangleq$ CHOOSE $b \in \mathcal{B} : b > c \land$ $\nexists\, m \in sent : m.type = \text{"1a"} \land m.bal = b$ $Preempt(p \in \mathcal{P}) \triangleq \exists\, m \in sent :$ $\land\, m.type = \text{"preempt"} \land m.to = p$ $\land\, m.bal > proBallot[p]$ $\land\, proBallot' = [proBallot$ EXCEPT $![p] =$ $\quad NewBallot(m.bal)]$ $\land$ UNCHANGED $\langle sent, aVoted, aBal \rangle$ $Phase1a(p \in \mathcal{P}) \triangleq$ $\land\, \nexists\, m \in sent : (m.type = \text{"1a"}) \land$ $\quad (m.bal = proBallot[p])$ $\land\, Send([type \mapsto \text{"1a"},$ $\quad from \mapsto p, bal \mapsto proBallot[p]])$ $\land$ UNCHANGED $\langle aVoted, aBal, proBallot \rangle$	$Phase1a(p \in \mathcal{P}) \triangleq \exists\, b \in \mathcal{B} :$ $\land \lor \exists\, m \in sent :$ $\quad \land\, m.type = \text{"preempt"} \land m.to = p$ $\quad \land\, Receive(m)$ $\quad \land \forall\, m2 \in received' : m2.to = p \land$ $\quad\quad m2.type = \text{"preempt"} \Rightarrow b > m2.bal$ $\lor \land \nexists\, m \in sent : m.type = \text{"1a"} \land$ $\quad\quad m.from = p$ $\quad \land$ UNCHANGED $\langle received \rangle$ $\land\, Send([type \mapsto \text{"1a"}, from \mapsto p, bal \mapsto b])$

Figure 7: Preemption in Multi-Paxos

important player in this decrease was the removal of operator *MaxVotedBallotInSlot* from Chand et al.'s specifications. This operator was defined as

$$MaxVotedBallotInSlot(D, s) \triangleq \mathsf{max}(\{d.bal : d \in \{d \in D : d.slot = s\}\})$$

Five lemmas were needed in Chand et al.'s proof to assert basic properties of the operator. For example, lemma *MVBISType* stated that if $D \subseteq [bal : \mathcal{B}, slot : \mathcal{S}, val : \mathcal{V}]$, then the result of the operator is in $\mathcal{B} \cup \{-1\}$. Removing these lemmas and their proofs alone resulted in a decrease of about 100 lines (about 10%) in proof size.

5 Results

Table 1 summarizes the results of our specifications and proofs that use only message history variables, compared with those by Lamport et al. and Chand et al. We observe an improvement of around 25% across all stats for Basic Paxos and a staggering 50% for Multi-Paxos and Multi-Paxos with Preemption. Following, we list some important results:

- The specification size decreased by 13 lines (25%) for Basic Paxos, from 52 lines for Lamport et al.'s specification to 39 lines for ours. For Multi-Paxos, the decrease is 36 lines (46%), from 78 lines for Chand et al.'s to 42 lines for ours, and for Multi-Paxos with Preemption, the decrease is 45 lines (46%), from 97 to 52.

- The total number of manually written invariants decreased by 54% overall—by 9 (60%) from 15 to 6 for Basic Paxos, by 8 (50%) from 16 to 8 for Multi-Paxos, and by 9 (53%) from 17 to 8 for Multi-Paxos with Preemption. This drastic decrease is because we do not maintain the variables *maxBal*, *maxVBal*, and *maxVal* as explained in Section 3.

- The proof size for Basic Paxos decreased by 83 lines (27%), from 310 to 227. This decrease is attributed to the fact that our specification does not use other state variables

Metric	Basic Paxos			Multi-Paxos			Multi-Paxos with Preemption		
	Lam	Us	Decr	Cha	Us	Decr	Cha	Us	Decr
Specification size	52	39, 33*	25%	78	42	46%	97	52	46%
# invariants	15	6	60%	16	8	50%	17	8	53%
# type invariants	4	1	75%	4	1	75%	5	1	80%
# process invariants	4	0	100%	4	0	100%	4	0	100%
# message invariants	7	5	29%	8	7	13%	8	7	13%
Proof size	310	227, 115*	27%	988	520	47%	1032	538	48%
Type invariants' proof size	22	21, 12*	5%	54	34	37%	75	38	49%
Process invariants' proof size	27	0, 0*	100%	136	0	100%	141	0	100%
1b† invariants' proof size	21	15, 9*	29%	133	70	47%	133	70	47%
2a† invariants' proof size	73	57, 21*	22%	264	120	55%	269	120	55%
2b† invariants' proof size	14	12, 7*	14%	94	73	22%	94	73	22%
# proofs by contradiction	2	0	100%	3	0	100%	3	0	100%
# obligations in TLAPS	239	182	24%	918	468	49%	959	491	49%
Type inv proof obligations	17	17	0%	69	52	25%	100	60	40%
Process inv proof obligations	39	0	100%	163	0	100%	173	0	100%
1b† inv proof obligations	12	10	17%	160	80	50%	160	80	50%
2a† inv proof obligations	62	52	16%	241	145	40%	249	145	42%
2b† inv proof obligations	9	9	0%	77	44	43%	77	44	43%
TLAPS check time (seconds)	42	31	26%	>191**	80	>58%	>208**	90	>57%

Table 1: Summary of results. Lam is for Lamport et al. [22], Cha is from Chand et al. [3], Us is ours in this paper, and Decr is percentage of decrease by Us from Lam and Cha.
Specification and proof sizes are measured in lines excluding comments and empty lines.
* indicates a number for the specification and proof in Appendices A and C respectively, after removing unnecessary line breaks from default latex generated by TLA$^+$ Tools.
† 1b invariants are (I10)–(I12), 2a invariants are (I13) and (I14), and 2b invariants are (I9) and (I15) for Basic Paxos in Figure 6, and corresponding ones for Multi-Paxos and Multi-Paxos with Preemption in [3].
An obligation is a condition that TLAPS checks.
Check time is on an Intel i7-4720HQ 2.6 GHz CPU with 16 GB of memory, running 64-bit Windows 10 Home (v1709 b16299.98) and TLAPS 1.5.4.
** indicates that TLAPS 1.5.4 failed to check and gave up after that number of seconds.

besides *sent*. This decrease is 468 lines (47%), from 988 to 520, for Multi-Paxos, and is 494 lines (48%), from 1032 to 538 for Multi-Paxos with Preemption.

- Proof by contradiction is used twice in the proof by Lamport et al. and thrice for the proofs in Chand et al. We were able to remove all of these because our specification uses queries as opposed to derived variables. The motive behind removing proofs by contradiction is to have easier to understand constructive proofs.

- The total number of proof obligations decreased by 46% overall—by 57 (24%) from 239 to 182 for Basic Paxos, by 450 (49%) from 918 to 468 for Multi-Paxos, and by 468 (49%) from 959 to 491 for Multi-Paxos with Preemption.

- The proof-checking time decreased by 11 seconds (26%), from 42 to 31 for Basic Paxos. For Multi-Paxos and Multi-Paxos with Preemption, TLAPS took over 3 minutes for the proofs in [3] and failed (due to updates in the new version of TLAPS) to check the proofs of about 5 obligations. In contrast, our proofs were able to be checked completely in 1.5 minutes or less.

6 Related work and conclusion

History variables. History variables have been at the center of much debate since they were introduced in the early 1970s [7, 6, 8]. Owicki and Gries [28] use them in an effort to prove properties of parallel programs, criticized by Lamport in his writings [14]. Contrary to ours, their history variables were ghost or auxiliary variables introduced for the sole purpose of simpler proofs. Our history variables are *sent* and *received*, whose contents are actually processed in all distributed system implementations.

Recently, Lamport and Merz [21] present rules to add history variables, among other auxiliary variables, to a low-level specification so that a refinement mapping from a high-level one can be established. The idea is to prove invariants in the high-level specification that serves as an abstraction of the low-level specification. In contrast, we focus on high-level specifications because our target executable language is DistAlgo, and efficient lower-level implementations can be generated systematically from high-level code.

Specification and verification. A number of systems [33, 9, 5], models [35, 4, 27], and methods [29, 12, 13, 1] have been developed in the past to specify distributed algorithms and mechanically check proofs of the safety and liveness properties of the algorithms. This work is orthogonal to them in the sense that the idea of maintaining only message history variables can be incorporated in their specifications as well.

Closer to our work in terms of the specification is the work by Padon et al. [29], which does not define any variable and instead defines predicate relations which would correspond to manipulations of our history variables. For example, $Send([type \mapsto \text{“1a”}, bal \mapsto b])$ is denoted by $start\_round\_msg(b)$. Instead of using TLA$^+$, the temporal logic of actions, they specify Paxos in first-order logic to later exploit benefits of Effectively Propositional Logic, such as satisfiability being decidable in it.

In contrast, we present a method to specify distributed algorithms using history variables, implementable in high-level executable languages like DistAlgo, and then show (i) how such specifications require fewer invariants for proofs and (ii) how several important invariants can be systematically derived.

Conclusion. We have shown that using message history variables can lead to simpler specifications and easier proofs of challenging distributed algorithms. Future work includes applying our method in specification and proofs of other complex distributed algorithms, and extending our method for proving liveness properties.

Acknowledgements. This work was supported in part by National Science Foundation grants CCF-1414078 and CCF-1248184 and Office of Naval Research grant N000141512208. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of these agencies.

References

[1] Martín Abadi and Leslie Lamport. The existence of refinement mappings. *Theoretical Computer Science*, 82(2):253–284, 1991.

[2] Saksham Chand and Yanhong A Liu. Simpler specifications and easier proofs of distributed algorithms using history variables. In *NASA Formal Methods Symposium*, pages 70–86. Springer, 2018.

[3] Saksham Chand, Yanhong A Liu, and Scott D Stoller. Formal verification of multi-paxos for distributed consensus. In *FM 2016: Formal Methods: 21st International Symposium, Limassol, Cyprus, November 9-11, 2016, Proceedings 21*, pages 119–136. Springer, 2016.

[4] Bernadette Charron-Bost and André Schiper. The Heard-Of model: computing in distributed systems with benign faults. *Distributed Computing*, 22(1):49–71, 2009.

[5] Kaustuv Chaudhuri, Damien Doligez, Leslie Lamport, and Stephan Merz. The TLA$^+$ proof system: Building a heterogeneous verification platform. In *Proceedings of the 7th International colloquium conference on Theoretical aspects of computing*, pages 44–44. Springer-Verlag, 2010.

[6] Edmund M Clarke. Proving correctness of coroutines without history variables. *Acta Informatica*, 13(2):169–188, 1980.

[7] Maurice Clint. Program proving: coroutines. *Acta informatica*, 2(1):50–63, 1973.

[8] Maurice Clint. On the use of history variables. *Acta Informatica*, 16(1):15–30, 1981.

[9] Cezara Drăgoi, Thomas A Henzinger, and Damien Zufferey. PSync: a partially synchronous language for fault-tolerant distributed algorithms. *ACM SIGPLAN Notices*, 51(1):400–415, 2016.

[10] Mario Gerla, Eun-Kyu Lee, Giovanni Pau, and Uichin Lee. Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds. In *Internet of Things (WF-IoT), 2014 IEEE World Forum on*, pages 241–246. IEEE, 2014.

[11] Michael Gorbovitski. *A system for invariant-driven transformations*. PhD thesis, Computer Science Department, Stony Brook University, 2011.

[12] Chris Hawblitzel, Jon Howell, Manos Kapritsos, Jacob R Lorch, Bryan Parno, Michael L Roberts, Srinath Setty, and Brian Zill. IronFleet: proving practical distributed systems correct. In *Proceedings of the 25th Symposium on Operating Systems Principles*, pages 1–17. ACM, 2015.

[13] Philipp Küfner, Uwe Nestmann, and Christina Rickmann. Formal verification of distributed algorithms. *Theoretical Computer Science*, pages 209–224, 2012.

[14] Leslie Lamport. My writings :: Proving the correctness of multiprocess programs. https://lamport.azurewebsites.net/pubs/pubs.html. Accessed: 2017-10-10.

[15] Leslie Lamport. The implementation of reliable distributed multiprocess systems. *Computer Networks (1976)*, 2(2):95–114, 1978.

[16] Leslie Lamport. Time, clocks, and the ordering of events in a distributed system. *Communications of the ACM*, 21(7):558–565, 1978.

[17] Leslie Lamport. The temporal logic of actions. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 16(3):872–923, 1994.

[18] Leslie Lamport. The part-time parliament. *ACM Transactions on Computer Systems (TOCS)*, 16(2):133–169, 1998.

[19] Leslie Lamport. Paxos made simple. *ACM Sigact News*, 32(4):18–25, 2001.

[20] Leslie Lamport. How to write a 21st century proof. *Journal of fixed point theory and applications*, 11(1):43–63, 2012.

[21] Leslie Lamport and Stephan Merz. Auxiliary variables in TLA$^+$. *ArXiv e-prints*, March 2017.

[22] Leslie Lamport, Stephan Merz, and Damien Doligez. Paxos.tla. https://github.com/tlaplus/v1-tlapm/blob/master/examples/paxos/Paxos.tla. Last modified Fri Nov 28 10:39:17 PST 2014 by Lamport. Accessed Feb 6, 2018.

[23] Yanhong A Liu, , Jon Brandvein, Scott D Stoller, and Bo Lin. Demand-driven incremental object queries. In *Proceedings of the 18th International Symposium on Principles and Practice of Declarative Programming (PPDP)*, pages 228–241. ACM, 2016.

[24] Yanhong A Liu. *Systematic Program Design: From Clarity To Efficiency*. Cambridge University Press, 2013.

[25] Yanhong A Liu, Scott D Stoller, and Bo Lin. From clarity to efficiency for distributed algorithms. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 39(3):12, 2017.

[26] Yanhong A Liu, Scott D Stoller, Bo Lin, and Michael Gorbovitski. From clarity to efficiency for distributed algorithms. In *Proceedings of the 27th ACM SIGPLAN Conference on Object-Oriented Programming, Systems, Languages and Applications (OOPSLA)*, pages 395–410. ACM, 2012.

[27] Nancy A Lynch and Mark R Tuttle. Hierarchical correctness proofs for distributed algorithms. In *Proceedings of the sixth annual ACM Symposium on Principles of distributed computing*, pages 137–151. ACM, 1987.

[28] Susan Owicki and David Gries. An axiomatic proof technique for parallel programs i. *Acta informatica*, 6(4):319–340, 1976.

[29] Oded Padon, Giuliano Losa, Mooly Sagiv, and Sharon Shoham. Paxos made EPR: decidable reasoning about distributed protocols. *Proceedings of the ACM on Programming Languages*, 1(OOPSLA):108, 2017.

[30] Robert Paige and Shaye Koenig. Finite differencing of computable expressions. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 4(3):402–454, 1982.

[31] Tom Rothamel and Yanhong A Liu. Generating incremental implementations of object-set queries. In *Proceedings of the 7th international conference on Generative programming and component engineering*, pages 55–66. ACM, 2008.

[32] Klaus Schilling. Perspectives for miniaturized, distributed, networked cooperating systems for space exploration. *Robotics and Autonomous Systems*, 90:118–124, 2017.

[33] Ilya Sergey, James R Wilcox, and Zachary Tatlock. Programming and proving with distributed protocols. *Proceedings of the ACM on Programming Languages*, 2(POPL):28, 2017.

[34] Florian Tschorsch and Björn Scheuermann. Bitcoin and beyond: A technical survey on decentralized digital currencies. *IEEE Communications Surveys & Tutorials*, 18(3):2084–2123, 2016.

[35] James R Wilcox, Doug Woos, Pavel Panchekha, Zachary Tatlock, Xi Wang, Michael D Ernst, and Thomas Anderson. Verdi: a framework for implementing and formally verifying distributed systems. In *ACM SIGPLAN Notices*, volume 50, pages 357–368. ACM, 2015.

[36] Pamela Zave. Using lightweight modeling to understand Chord. *ACM SIGCOMM Computer Communication Review*, 42(2):49–57, 2012.

A TLA$^+$ specification of Basic Paxos

MODULE *PaxosSpec*

This is a specification in TLA$^+$ of *BasicPaxos*.

EXTENDS *Integers*, *TLAPS*, *NaturalsInduction*

CONSTANTS $\mathcal{A}$, $\mathcal{Q}$, $\mathcal{V}$ Sets of acceptors, quorums of acceptors, and values to propose

VARIABLES *sent* Set of sent messages

ASSUME $QuorumAssumption \triangleq \mathcal{Q} \subseteq \text{SUBSET } \mathcal{A} \land \forall\, Q1, Q2 \in \mathcal{Q} : Q1 \cap Q2 \neq \emptyset$

$\mathcal{B} \triangleq \mathbb{N}$ Set of ballots
$vars \triangleq \langle sent \rangle$
$Send(m) \triangleq sent' = sent \cup \{m\}$
$None \triangleq \text{CHOOSE } v : v \notin \mathcal{V}$

Phase 1a: A 1a message with ballot b is sent by some proposer (to all processes).

$Phase1a(b) \triangleq Send([type \mapsto \text{“1a”}, bal \mapsto b])$

Phase 1b: For an acceptor a, if there is a 1a message m with ballot $m.bal$ that is higher than the highest it has seen, a sends a 1b message with $m.bal$ alongwith the highest-numbered pair it has voted for.

$2bs(a) \triangleq \{m \in sent : m.type = \text{“2b”} \land m.acc = a\}$
$max\_prop(a) \triangleq \text{IF } 2bs(a) = \emptyset \text{ THEN } \{[bal \mapsto -1, val \mapsto None]\}$
$\qquad\qquad\qquad\qquad\qquad\qquad \text{ELSE } \{m \in 2bs : \forall\, m2 \in 2bs : m.bal \geq m2.bal\}$

$Phase1b(a) \triangleq \exists\, m \in sent, r \in max\_prop(a) :$
$\quad \land\, m.type = \text{“1a”}$
$\quad \land\, \forall\, m2 \in sent : m2.type \in \{\text{“1b”}, \text{“2b”}\} \land m2.acc = a \Rightarrow m.bal > m2.bal$
$\quad \land\, Send([type \mapsto \text{“1b”}, bal \mapsto m.bal, maxVBal \mapsto r.bal, maxVal \mapsto r.val, acc \mapsto a])$

Phase 2a: If there is no 2a message in *sent* with ballot b, and a quorum of acceptors has sent a set S of 1b messages with ballot b, a proposer sends a 2a message with ballot b and value v, where v is the value with the highest ballot in S, or is any value in $\mathcal{V}$ if no acceptor that responded in S has voted for anything.

$Phase2a(b) \triangleq$
$\quad \land\, \nexists\, m \in sent : (m.type = \text{“2a”}) \land (m.bal = b)$
$\quad \land\, \exists\, v \in \mathcal{V},\ Q \in \mathcal{Q},\ S \subseteq \{m \in sent : m.type = \text{“1b”} \land m.bal = b\} :$
$\qquad \land\, \forall\, a \in Q : \exists\, m \in S : m.acc = a$
$\qquad \land\ \lor\, \forall\, m \in S : m.maxVBal = -1$
$\qquad\quad \lor\, \exists\, b2 \in 0\,..\,(b-1) :$
$\qquad\qquad \land\, \forall\, m \in S : m.maxVBal \leq b2$
$\qquad\qquad \land\, \exists\, m \in S : m.maxVBal = b2 \land m.maxVal = v$
$\qquad \land\, Send([type \mapsto \text{“2a”}, bal \mapsto b, val \mapsto v])$

Phase 2b: For an acceptor a, if there is a 2a message m with ballot $m.bal$ that is higher than or equal to the highest it has seen, a sends a 2b message with $m.bal$ and $m.val$.

$Phase2b(a) \triangleq \exists\, m \in sent :$
$\quad \land\, m.type = \text{“2a”}$
$\quad \land\, \forall\, m2 \in sent : m2.type \in \{\text{“1b”}, \text{“2b”}\} \land m2.acc = a \Rightarrow m.bal \geq m2.bal$
$\quad \land\, Send([type \mapsto \text{“2b”}, bal \mapsto m.bal, val \mapsto m.val, acc \mapsto a])$

$Init \triangleq sent = \emptyset$
$Next \triangleq \lor\, \exists\, b \in \mathcal{B} : Phase1a(b) \lor Phase2a(b)$
$\qquad\ \lor\, \exists\, a \in \mathcal{A} : Phase1b(a) \lor Phase2b(a)$
$Spec \triangleq Init \land \Box[Next]_{vars}$

B Safety property to prove for Basic Paxos and invariants used in proof

MODULE *PaxosProp*

VotedForIn(a, v, b) means that acceptor a has sent some 2b message m with $m.val$ equal to v and $m.bal$ equal to b. This specifies that acceptor a has voted the pair $\langle v, b \rangle$.

$$VotedForIn(a,\ v,\ b) \triangleq \exists\, m \in sent : m.type = \text{“2b”} \land m.val = v \land m.bal = b \land m.acc = a$$

ChosenIn(v, b) means that every acceptor in some quorum Q has voted the pair $\langle v, b \rangle$.

$$ChosenIn(v,\ b) \triangleq \exists\, Q \in \mathcal{Q} : \forall\, a \in Q : VotedForIn(a,\ v,\ b)$$

Chosen(v) means that for some ballot b, *ChosenIn*(v, b) holds.

$$Chosen(v) \triangleq \exists\, b \in \mathcal{B} : ChosenIn(v,\ b)$$

WontVoteIn(a, b) means that acceptor a has seen a higher ballot than b, and did not and will not vote any value with ballot b.

$$\begin{aligned} WontVoteIn(a,\ b) \triangleq\ & \land \forall\, v \in \mathcal{V} : \neg VotedForIn(a,\ v,\ b) \\ & \land \exists\, m \in sent : m.type \in \{\text{“1b”}, \text{“2b”}\} \land m.acc = a \land m.bal > b \end{aligned}$$

SafeAt(v, b) means that no value except perhaps v has been or will be chosen in any ballot lower than b.

$$SafeAt(v,\ b) \triangleq \forall\, b2 \in 0\,..\,(b-1) : \exists\, Q \in \mathcal{Q} : \forall\, a \in Q : VotedForIn(a,\ v,\ b2) \lor WontVoteIn(a,\ b2)$$

Messages defines the set of valid messages. *TypeOK* defines invariants for the types of the variables.

$$\begin{aligned} Messages \triangleq\ & [type : \{\text{“1a”}\},\ bal : \mathcal{B}] \cup \\ & [type : \{\text{“1b”}\},\ bal : \mathcal{B},\ maxVBal : \mathcal{B} \cup \{-1\},\ maxVal : \mathcal{V} \cup \{None\},\ acc : \mathcal{A}] \cup \\ & [type : \{\text{“2a”}\},\ bal : \mathcal{B},\ val : \mathcal{V}] \cup \\ & [type : \{\text{“2b”}\},\ bal : \mathcal{B},\ val : \mathcal{V},\ acc : \mathcal{A}] \\ TypeOK \triangleq\ & sent \subseteq Messages \end{aligned}$$

MsgInv defines properties satisfied by the contents of messages, for 1b, 2a, and 2b messages.

$$\begin{aligned} MsgInv \triangleq\ & \forall\, m \in sent : \\ & \land\ m.type = \text{“1b”} \Rightarrow \land\ VotedForIn(m.acc,\ m.maxVal,\ m.maxVBal) \lor m.maxVBal = -1 \\ & \qquad\qquad \land \forall\, b \in m.maxVBal + 1\,..\,m.bal - 1 : \nexists\, v \in \mathcal{V} : VotedForIn(m.acc,\ v,\ b) \\ & \land\ m.type = \text{“2a”} \Rightarrow \land\ SafeAt(m.val,\ m.bal) \\ & \qquad\qquad \land \forall\, m2 \in sent : (m2.type = \text{“2a”} \land m2.bal = m.bal) \Rightarrow m2 = m \\ & \land\ m.type = \text{“2b”} \Rightarrow \exists\, m2 \in sent : m2.type = \text{“2a”} \land m2.bal = m.bal \land m2.val = m.val \end{aligned}$$

Inv is the complete inductive invariant.

$$Inv \triangleq TypeOK \land MsgInv$$

Safe states that at most one value can be chosen.

$$Safe \triangleq \forall\, v1,\ v2 \in \mathcal{V} : Chosen(v1) \land Chosen(v2) \Rightarrow v1 = v2$$

C TLAPS checked proof of Basic Paxos

MODULE *PaxosProof*

The following two lemmas are straightforward consequences of the predicates defined above.

LEMMA $VotedInv \triangleq MsgInv \land TypeOK \Rightarrow \forall a \in \mathcal{A}, v \in \mathcal{V}, b \in \mathcal{B} : VotedForIn(a, v, b) \Rightarrow SafeAt(v, b)$
BY DEF *VotedForIn*, *MsgInv*, *Messages*, *TypeOK*

LEMMA $VotedOnce \triangleq MsgInv \Rightarrow \forall a1, a2 \in \mathcal{A}, v1, v2 \in \mathcal{V}, b \in \mathcal{B} :$
$VotedForIn(a1, v1, b) \land VotedForIn(a2, v2, b) \Rightarrow v1 = v2$
BY DEF *MsgInv*, *VotedForIn*

Lemma *SafeAtStable* asserts that if $SafeAt(v, b)$ holds then $SafeAt(v, b)$ holds in the next state as well.

LEMMA $SafeAtStable \triangleq Inv \land Next \Rightarrow \forall v \in \mathcal{V}, b \in \mathcal{B} : SafeAt(v, b) \Rightarrow SafeAt(v, b)'$
⟨1⟩. SUFFICES ASSUME *Inv*, *Next*, NEW $v \in \mathcal{V}$, NEW $b \in \mathcal{B}$, $SafeAt(v, b)$
PROVE $SafeAt(v, b)'$ OBVIOUS
⟨1⟩. USE DEF *Send*, *Inv*, $\mathcal{B}$
⟨1⟩1. ASSUME NEW $b2 \in \mathcal{B}$, $Phase1a(b2)$ PROVE $SafeAt(v, b)'$
BY ⟨1⟩1, SMT DEF *SafeAt*, *Phase1a*, *VotedForIn*, *WontVoteIn*
⟨1⟩2. ASSUME NEW $a \in \mathcal{A}$, $Phase1b(a)$ PROVE $SafeAt(v, b)'$
BY ⟨1⟩2, *QuorumAssumption*, SMTT(60) DEF *TypeOK*, *SafeAt*, *WontVoteIn*, *VotedForIn*, *Phase1b*
⟨1⟩3. ASSUME NEW $b2 \in \mathcal{B}$, $Phase2a(b2)$ PROVE $SafeAt(v, b)'$
BY ⟨1⟩3, *QuorumAssumption*, SMT DEF *TypeOK*, *SafeAt*, *WontVoteIn*, *VotedForIn*, *Phase2a*
⟨1⟩4. ASSUME NEW $a \in \mathcal{A}$, $Phase2b(a)$ PROVE $SafeAt(v, b)'$
⟨2⟩1. PICK $m \in sent : Phase2b(a)!(m)$ BY ⟨1⟩4 DEF *Phase2b*
⟨2⟩2. $\forall a2 \in \mathcal{A}, b2 \in \mathcal{B}, v2 \in \mathcal{V} : VotedForIn(a2, v2, b2) \Rightarrow VotedForIn(a2, v2, b2)'$
BY ⟨2⟩1 DEF *TypeOK*, *VotedForIn*
⟨2⟩3. ASSUME NEW $a2 \in \mathcal{A}$, NEW $b2 \in \mathcal{B}$, $WontVoteIn(a2, b2)$, NEW $v2 \in \mathcal{V}$
PROVE $\neg VotedForIn(a2, v2, b2)'$
⟨3⟩1. PICK $m1 \in sent : m1.type \in \{$"1b", "2b"$\} \land m1.acc = a2 \land m1.bal > b2$ BY ⟨2⟩3 DEF *WontVoteIn*
⟨3⟩2. $a2 = a \Rightarrow b2 \neq m.bal$ BY ⟨2⟩1, ⟨2⟩3, ⟨3⟩1, $a2 = a \Rightarrow m.bal \geq m1.bal$ DEF *TypeOK*, *Messages*
⟨3⟩3. $a2 \neq a \Rightarrow \neg VotedForIn(a2, v2, b2)'$ BY ⟨2⟩1, ⟨2⟩3 DEF *WontVoteIn*, *VotedForIn*
⟨3⟩. QED BY ⟨2⟩1, ⟨2⟩2, ⟨2⟩3, ⟨3⟩2, ⟨3⟩3 DEF *Phase2b*, *VotedForIn*, *WontVoteIn*, *TypeOK*, *Messages*, *Send*
⟨2⟩4. $\forall a2 \in \mathcal{A}, b2 \in \mathcal{B} : WontVoteIn(a2, b2) \Rightarrow WontVoteIn(a2, b2)'$ BY ⟨2⟩1, ⟨2⟩3 DEF *WontVoteIn*, *Send*
⟨2⟩. QED BY ⟨2⟩2, ⟨2⟩4, *QuorumAssumption* DEF *SafeAt*
⟨1⟩. QED BY ⟨1⟩1, ⟨1⟩2, ⟨1⟩3, ⟨1⟩4 DEF *Next*

Invariant asserts the temporal formula that if *Spec* holds then *Inv* always holds.

THEOREM $Invariant \triangleq Spec \Rightarrow \Box Inv$
⟨1⟩. USE DEF $\mathcal{B}$, *2bs*, *max_prop*
⟨1⟩1. $Init \Rightarrow Inv$ BY DEF *Init*, *Inv*, *TypeOK*, *MsgInv*, *VotedForIn*
⟨1⟩2. $Inv \land [Next]_{vars} \Rightarrow Inv'$
⟨2⟩. SUFFICES ASSUME *Inv*, *Next* PROVE Inv'
BY DEF *vars*, *Inv*, *TypeOK*, *MsgInv*, *VotedForIn*, *SafeAt*, *WontVoteIn*
⟨2⟩. USE DEF *Inv*

⟨2⟩1 proves $TypeOK'$ for *Next*. Each of ⟨3⟩1-4 assumes the action of a phase and proves $TypeOK'$ for that case.

⟨2⟩1. $TypeOK'$
⟨3⟩1. ASSUME NEW $b \in \mathcal{B}$, $Phase1a(b)$ PROVE $TypeOK'$ BY ⟨3⟩1 DEF *TypeOK*, *Phase1a*, *Send*, *Messages*
⟨3⟩2. ASSUME NEW $b \in \mathcal{B}$, $Phase2a(b)$ PROVE $TypeOK'$
⟨4⟩. PICK $v \in \mathcal{V} : Send([type \mapsto$ "2a", $bal \mapsto b, val \mapsto v])$ BY ⟨3⟩2 DEF *Phase2a*
⟨4⟩. QED BY DEF *TypeOK*, *Send*, *Messages*
⟨3⟩3. ASSUME NEW $a \in \mathcal{A}$, $Phase1b(a)$ PROVE $TypeOK'$
⟨4⟩. PICK $m \in sent, r \in max\_prop(a) : Phase1b(a)!(m, r)$ BY ⟨3⟩3 DEF *Phase1b*
⟨4⟩. QED BY DEF *Send*, *TypeOK*, *Messages*
⟨3⟩4. ASSUME NEW $a \in \mathcal{A}$, $Phase2b(a)$ PROVE $TypeOK'$
⟨4⟩. PICK $m \in sent : Phase2b(a)!(m)$ BY ⟨3⟩4 DEF *Phase2b*
⟨4⟩. QED BY DEF *Send*, *TypeOK*, *Messages*
⟨3⟩. QED BY ⟨3⟩1, ⟨3⟩2, ⟨3⟩3, ⟨3⟩4 DEF *Next*

⟨2⟩2 proves $MsgInv'$ for *Next*. Each of ⟨3⟩1-4 assumes the action of a phase and proves $MsgInv'$ for that case.

⟨2⟩2. $MsgInv'$
⟨3⟩1. ASSUME NEW $b \in \mathcal{B}$, $Phase1a(b)$ PROVE $MsgInv'$
⟨4⟩1. $\forall a2, v2, b2 : VotedForIn(a2, v2, b2)' \equiv VotedForIn(a2, v2, b2)$ BY ⟨3⟩1 DEF *Send*, *VotedForIn*, *Phase1a*
⟨4⟩. QED BY ⟨3⟩1, ⟨4⟩1, *QuorumAssumption*, *SafeAtStable* DEF *Phase1a*, *MsgInv*, *TypeOK*, *Messages*, *Send*
⟨3⟩2. ASSUME NEW $a \in \mathcal{A}$, $Phase1b(a)$ PROVE $MsgInv'$
⟨4⟩. PICK $m \in sent, r \in max\_prop(a) : Phase1b(a)!(m, r)$ BY ⟨3⟩2 DEF *Phase1b*

⟨4⟩. DEFINE $m2 \triangleq [type \mapsto$ "1b", $bal \mapsto m.bal$, $maxVBal \mapsto r.bal$, $maxVal \mapsto r.val$, $acc \mapsto a]$
⟨4⟩1. $\forall\, a2, v2, b2 : VotedForIn(a2, v2, b2)' \equiv VotedForIn(a2, v2, b2)$ BY DEF *Send*, *VotedForIn*
⟨4⟩2. $VotedForIn(m2.acc, m2.maxVal, m2.maxVBal) \lor m2.maxVBal = -1$
BY DEF *TypeOK*, *Messages*, *VotedForIn*
⟨4⟩3. $\forall\, b \in (r.bal + 1) \,..\, (m2.bal - 1) : \nexists\, v \in \mathcal{V} : VotedForIn(m2.acc, v, b)$
BY DEF *TypeOK*, *Messages*, *VotedForIn*, *Send*
⟨4⟩. QED BY ⟨4⟩1, ⟨4⟩2, ⟨4⟩3, *SafeAtStable* DEF *MsgInv*, *TypeOK*, *Messages*, *Send*
⟨3⟩3. ASSUME NEW $b \in \mathcal{B}$, $Phase2a(b)$ PROVE $MsgInv'$
⟨4⟩1. $\nexists\, m \in sent : (m.type =$ "2a") $\land (m.bal = b)$ BY ⟨3⟩3 DEF *Phase2a*
⟨4⟩2. PICK $v \in \mathcal{V}$, $Q \in \mathcal{Q}$, $S \subseteq \{m \in sent : m.type =$ "1b" $\land\, m.bal = b\}$:
$\land\, \forall\, a \in Q : \exists\, m \in S : m.acc = a$
$\land\, \lor\, \forall\, m \in S : m.maxVBal = -1$
$\lor\, \exists\, b2 \in 0 \,..\, (b-1) :$
$\land\, \forall\, m \in S : m.maxVBal \leq b2$
$\land\, \exists\, m \in S : m.maxVBal = b2 \land m.maxVal = v$
$\land\, Send([type \mapsto$ "2a", $bal \mapsto b$, $val \mapsto v])$ BY ⟨3⟩3 DEF *Phase2a*
⟨4⟩3. $sent' = sent \cup \{[type \mapsto$ "2a", $bal \mapsto b$, $val \mapsto v]\}$ BY ⟨4⟩2 DEF *Send*
⟨4⟩4. $\forall\, a2, v2, b2 : VotedForIn(a2, v2, b2)' \equiv VotedForIn(a2, v2, b2)$ BY ⟨4⟩3 DEF *VotedForIn*
⟨4⟩5. $\forall\, m, m2 \in sent' : m.type =$ "2a" $\land\, m2.type =$ "2a" $\land\, m2.bal = m.bal \Rightarrow m2 = m$
BY ⟨4⟩1, ⟨4⟩3, Isa DEF *MsgInv*
⟨4⟩6. $SafeAt(v, b)$
⟨5⟩1. CASE $\forall\, m \in S : m.maxVBal = -1$ BY ⟨4⟩2, ⟨5⟩1 DEF *TypeOK*, *MsgInv*, *SafeAt*, *WontVoteIn*
⟨5⟩2. ASSUME NEW $b2 \in 0 \,..\, (b-1)$, $\forall\, m \in S : m.maxVBal \leq b2$,
NEW $m2 \in S$, $m2.maxVBal = b2$, $m2.maxVal = v$ PROVE $SafeAt(v, b)$
⟨6⟩. SUFFICES ASSUME NEW $b3 \in 0 \,..\, (b-1)$
PROVE $\exists\, Q2 \in \mathcal{Q} : \forall\, a \in Q2 : VotedForIn(a, v, b3) \lor WontVoteIn(a, b3)$ BY DEF *SafeAt*
⟨6⟩1. CASE $b3 \in 0 \,..\, (b2-1)$ BY ⟨5⟩2, ⟨6⟩1, *VotedInv* DEF *SafeAt*, *MsgInv*, *TypeOK*, *Messages*
⟨6⟩2. CASE $b3 = b2$
⟨7⟩1. $VotedForIn(m2.acc, v, b2)$ BY ⟨5⟩2 DEF *MsgInv*
⟨7⟩2. $\forall\, a2 \in Q, w \in \mathcal{V} : VotedForIn(a2, w, b2) \Rightarrow w = v$
BY ⟨7⟩1, *VotedOnce*, *QuorumAssumption* DEF *TypeOK*, *Messages*
⟨7⟩. QED BY ⟨4⟩2, ⟨6⟩2, ⟨7⟩2 DEF *WontVoteIn*
⟨6⟩3. CASE $b3 \in (b2+1) \,..\, (b-1)$ BY ⟨4⟩2, ⟨5⟩2, ⟨6⟩3 DEF *MsgInv*, *TypeOK*, *Messages*, *WontVoteIn*
⟨6⟩. QED BY ⟨6⟩1, ⟨6⟩2, ⟨6⟩3
⟨5⟩. QED BY ⟨4⟩2, ⟨5⟩1, ⟨5⟩2
⟨4⟩7. $(\forall\, m \in sent : m.type =$ "2a" $\Rightarrow SafeAt(m.val, m.bal))'$
BY ⟨4⟩3, ⟨4⟩6, *SafeAtStable* DEF *MsgInv*, *TypeOK*, *Messages*
⟨4⟩. QED BY ⟨4⟩3, ⟨4⟩4, ⟨4⟩5, ⟨4⟩7, $\forall\, m \in sent' \setminus sent : m.type \neq$ "1b" DEF *MsgInv*, *TypeOK*, *Messages*
⟨3⟩4. ASSUME NEW $a \in \mathcal{A}$, $Phase2b(a)$ PROVE $MsgInv'$
⟨4⟩. PICK $m \in sent : Phase2b(a)!(m)$ BY ⟨3⟩4 DEF *Phase2b*
⟨4⟩1. $\forall\, a2, v2, b2 : VotedForIn(a2, v2, b2) \Rightarrow VotedForIn(a2, v2, b2)'$ BY DEF *VotedForIn*, *Send*
⟨4⟩2. $\forall\, m2 \in sent : m2.type =$ "1b" $\Rightarrow \forall\, v \in \mathcal{V}, b2 \in (m2.maxVBal + 1) \,..\, (m2.bal - 1) :$
$\neg\, VotedForIn(m2.acc, v, b2) \Rightarrow \neg\, VotedForIn(m2.acc, v, b2)'$
BY DEF *Send*, *VotedForIn*, *MsgInv*, *TypeOK*, *Messages*
⟨4⟩. QED BY ⟨2⟩1, ⟨4⟩1, ⟨4⟩2, *SafeAtStable* DEF *MsgInv*, *Send*, *TypeOK*, *Messages*
⟨3⟩. QED BY ⟨3⟩1, ⟨3⟩2, ⟨3⟩3, ⟨3⟩4 DEF *Next*
⟨2⟩. QED BY ⟨2⟩1, ⟨2⟩2 DEF *Inv*
⟨1⟩. QED BY ⟨1⟩1, ⟨1⟩2, PTL DEF *Spec*

Safety asserts that *Spec* implies that *Safe* always holds.

THEOREM $Safety \triangleq Spec \Rightarrow \Box Safe$
⟨1⟩. USE DEF $\mathcal{B}$
⟨1⟩1. $Inv \Rightarrow Safe$
⟨2⟩. SUFFICES ASSUME *Inv*, NEW $v1 \in \mathcal{V}$, NEW $v2 \in \mathcal{V}$, NEW $b1 \in \mathcal{B}$, NEW $b2 \in \mathcal{B}$,
$ChosenIn(v1, b1)$, $ChosenIn(v2, b2)$, $b1 \leq b2$
PROVE $v1 = v2$ BY DEF *Safe*, *Chosen*
⟨2⟩1. CASE $b1 = b2$ BY ⟨2⟩1, *VotedOnce*, *QuorumAssumption*, SMTT(100) DEF *ChosenIn*, *Inv*
⟨2⟩2. CASE $b1 < b2$
⟨3⟩1. $SafeAt(v2, b2)$ BY *VotedInv*, *QuorumAssumption* DEF *ChosenIn*, *Inv*
⟨3⟩2. PICK $Q1 \in \mathcal{Q} : \forall\, a \in Q1 : VotedForIn(a, v2, b1) \lor WontVoteIn(a, b1)$ BY ⟨2⟩2, ⟨3⟩1 DEF *SafeAt*
⟨3⟩3. PICK $Q2 \in \mathcal{Q} : \forall\, a \in Q2 : VotedForIn(a, v1, b1)$ BY DEF *ChosenIn*
⟨3⟩4. QED BY ⟨3⟩2, ⟨3⟩3, *QuorumAssumption*, *VotedOnce*, Z3 DEF *WontVoteIn*, *Inv*
⟨2⟩. QED BY ⟨2⟩1, ⟨2⟩2
⟨1⟩. QED BY ⟨1⟩1, *Invariant*, PTL