MODULE *PConProof*

This is a specification of a variant of the classic *Paxos* consensus algorithm described in

AUTHOR = "*Leslie Lamport*", *TITLE* = "The Part-Time Parliament", journal = *ACM* Transactions on Computing Systems,
volume = 16,
Number = 2, Month = may, Year = 1998, pages = "133–169"

This algorithm was also described without proof in *Brian Oki*'s *Ph.D.* thesis.

It describes the actions that can be performed by leaders, but does not introduce explicit leader processes. More precisely, the specification is written as if there were a separate leader for each ballot.

This variant of the classic *Paxos* algorithm is an abstraction of an algorithm that is used in

AUTHOR = "*Leslie Lamport* and Dahlia *Malkhi* and *Lidong Zhou* ", *TITLE* = "Vertical *Paxos* and Primary-Backup Replication", Conference = "Proceedings of *PODC* 2009",
editor = {*Srikanta Tirthapura and Lorenzo Alvisi*},
publisher = {*ACM*}, *YEAR* = 2009, *PAGES* = "312–313"

and in

Cheap paxos United States Patent 7249280 Inventors: *Lamport*, *Leslie B.*
Massa, *Michael T.*
Filing Date:06/18/2004

In the classic *Paxos* algorithm, the leader sends a phase $2a$ message for a ballot b and value v that instructs acceptors to vote for v in ballot b. In terms of implementing the voting algorithm of module *VoteProof*, that $2a$ message serves two functions:

- It asserts that value v is safe at ballot b, so the acceptor can vote for it without violating invariant *VInv*2
- It tells the acceptors which single safe value they can vote for in ballot b, so they can vote for that value without violating *VInv*3.

The variant of the algorithm we specify here introduces phase $1c$ messages that perform the first function. The phase $2a$ message serves only the first function, being sent only if a $1c$ message had been sent for the value.

This variant of the algorithm is useful when reconfiguration is performed by using different sets of acceptors for different ballots. The leader propagates knowledge of what values are safe at ballot b so that the acceptors in the current configuration are no longer needed to determine that information. If the ballot b leader determines that all values are safe at b, then it sends a $1c$ message for every value and sends a phase $2a$ message only when it has a value to propose. The presence of the $1c$ messages removes dependency on the acceptors of ballots numbered b or lower for progress. (If the leader determines that only a single value is safe at b, then it sends the $1c$ and $2a$ messages together.)

In the algorithm described here, we do not include reconfiguration. Therefore, the sending of a $1c$ message serves only as a precondition for the sending of a $2a$ message with that value.

Classic *Paxos* and its variants maintain consensus in the presence of omission faults–faults in which a process fails to perform some enabled action or a message that is sent fails to be received. The safety specification, which is given by the *PlusCal* code, does not require that any action need ever be performed. A process need not execute an enabled action. Receipt of a message is modeled by a process performing the action enabled by that message having been sent, so message loss is also represented by a process not performing an enabled action. Thus, failures are never mentioned in the description of the algorithm.

EXTENDS *Integers*, *TLAPS*

The constant parameters and the set Ballots are the same as in the voting algorithm.

CONSTANT *Value*, *Acceptor*, *Quorum*

$$\text{ASSUME } QA \triangleq \begin{array}{l} \land \forall Q \in Quorum : Q \subseteq Acceptor \\ \land \forall Q1, Q2 \in Quorum : Q1 \cap Q2 \neq \{\} \end{array}$$

$$Ballot \triangleq Nat$$

We are going to have a leader process for each ballot and an acceptor process for each acceptor. So we can use the ballot numbers and the acceptors themselves as the identifiers for these processes, we assume that the set of ballots and the set of acceptors are disjoint. For good measure, we also assume that -1 is not an acceptor, although that is probably not necessary.

$$\text{ASSUME } BallotAssump \triangleq (Ballot \cup \{-1\}) \cap Acceptor = \{\}$$

We define *None* to be an unspecified value that is not in the set *Value*.

$$None \triangleq \text{CHOOSE } v : v \notin Value$$

This is a message-passing algorithm, so we begin by defining the set Message of all possible messages. The messages are explained below with the actions that send them. A message m with $m.type =$ "1a" is called a $1a$ message, and similarly for the other message types.

$$\begin{array}{rl} Message \triangleq & [type : \{\text{“1a”}\}, bal : Ballot] \\ \cup & [type : \{\text{“1b”}\}, acc : Acceptor, bal : Ballot, \\ & \ mbal : Ballot \cup \{-1\}, mval : Value \cup \{None\}] \\ \cup & [type : \{\text{“1c”}\}, bal : Ballot, val : Value] \\ \cup & [type : \{\text{“2a”}\}, bal : Ballot, val : Value] \\ \cup & [type : \{\text{“2b”}\}, acc : Acceptor, bal : Ballot, val : Value] \end{array}$$

The algorithm is easiest to understand in terms of the set *msgs* of all messages that have ever been sent. A more accurate model would use one or more variables to represent the messages actually in transit, and it would include actions representing message loss and duplication as well as message receipt.

In the current spec, there is no need to model message loss explicitly. The safety part of the spec says only what messages may be received and does not assert that any message actually is received. Thus, there is no difference between a lost message and one that is never received. The liveness property of the spec will make it clear what messages must be received (and hence either not lost or successfully retransmitted if lost) to guarantee progress.

Another advantage of maintaining the set of all messages that have ever been sent is that it allows us to define the state function *votes* that implements the variable of the same name in the voting algorithm without having to introduce a history variable.

****

In addition to the variable *msgs*, the algorithm uses four variables whose values are arrays indexed by acceptor, where for any acceptor a :

$maxBal[a]$ The largest ballot number in which a has participated

$maxVBal[a]$ The largest ballot number in which a has voted, or -1 if it has never voted.

maxVVal[*a*] If a has voted, then this is the value it voted for in ballot *maxVBal*; otherwise it equals *None*.

As in the voting algorithm, an execution of the algorithm consists of an execution of zero or more ballots. Different ballots may be in progress concurrently, and ballots may not complete (and need not even start). A ballot b consists of the following actions (which need not all occur in the indicated order).

*Phase*1*a* : The leader sends a $1a$ message for ballot b

*Phase*1*b* : If $maxBal[a] < b$, an acceptor a responds to the $1a$ message by setting $maxBal[a]$ to b and sending a $1b$ message to the leader containing the values of $maxVBal[a]$ and $maxVVal[a]$.

*Phase*1*c* : When the leader has received ballot-b $1b$ messages from a quorum, it determines some set of values that are safe at b and sends $1c$ messages for them.

*Phase*2*a* : The leader sends a $2a$ message for some value for which it has already sent a ballot-b $1c$ message.

*Phase*2*b* : Upon receipt of the $2a$ message, if $maxBal[a] \leq b$, an acceptor a sets $maxBal[a]$ and $maxVBal[a]$ to b, sets $maxVVal[a]$ to the value in the $2a$ message, and votes for that value in ballot b by sending the appropriate $2b$ message.

Here is the *PlusCal* code for the algorithm, which we call *PCon*.

--algorithm *PCon*{
  variables $maxBal = [a \in Acceptor \mapsto -1]$,
  $maxVBal = [a \in Acceptor \mapsto -1]$,
  $maxVVal = [a \in Acceptor \mapsto None]$,
  $msgs = \{\}$
  define {

$$sentMsgs(t, b) \triangleq \{m \in msgs : (m.type = t) \land (m.bal = b)\}$$

We define *ShowsSafeAt* so that $ShowsSafeAt(Q, b, v)$ is true for a quorum Q iff *msgs* contain ballot-b $1b$ messages from the acceptors in Q showing that v is safe at b.

$$\begin{array}{l} ShowsSafeAt(Q, b, v) \triangleq \\ \quad \text{LET } Q1b \triangleq \{m \in sentMsgs(\text{"1b"}, b) : m.acc \in Q\} \\ \quad \text{IN } \quad \land \forall a \in Q : \exists m \in Q1b : m.acc = a \\ \qquad\qquad \land \lor \forall m \in Q1b : m.mbal = -1 \\ \qquad\qquad\quad \lor \exists m1c \in msgs : \\ \qquad\qquad\qquad \land m1c = [type \mapsto \text{"1c"}, bal \mapsto m1c.bal, val \mapsto v] \\ \qquad\qquad\qquad \land \forall m \in Q1b : \land m1c.bal \geq m.mbal \\ \qquad\qquad\qquad\qquad\qquad\qquad \land (m1c.bal = m.mbal) \Rightarrow (m.mval = v) \end{array}$$

}

The following two macros send a message and a set of messages, respectively. These macros are so simple that they're hardly worth introducing, but they do make the processes a little easier to read.

macro $SendMessage(m)\{msgs := msgs \cup \{m\}\}$
macro $SendSetOfMessages(S)\{msgs := msgs \cup S\}$

The Actions

As before, we describe each action as a macro.

The leader for process *self* can execute a $Phase1a()$ action, which sends the ballot *self* $1a$ message.

$$\textbf{macro } Phase1a()\{SendMessage([type \mapsto \text{“1a”},\ bal \mapsto self])\}$$

Acceptor *self* can perform a $Phase1b(b)$ action, which is enabled iff $b > maxBal[self]$. The action sets $maxBal[self]$ to b and sends a phase $1b$ message to the leader containing the values of $maxVBal[self]$ and $maxVVal[self]$.

$$\begin{array}{l}
\textbf{macro } Phase1b(b)\{ \\
\quad \textbf{when } (b > maxBal[self]) \land (sentMsgs(\text{“1a”},\ b) \neq \{\})\,\textbf{;} \\
\quad maxBal[self] := b\,\textbf{;} \\
\quad SendMessage([type \mapsto \text{“1b”},\ acc \mapsto self,\ bal \mapsto b, \\
\qquad\qquad\qquad\qquad mbal \mapsto maxVBal[self],\ mval \mapsto maxVVal[self]])\,\textbf{;} \\
\}
\end{array}$$

The ballot *self* leader can perform a $Phase1c(S)$ action, which sends a set S of $1c$ messages indicating that the value in the *val* field of each of them is safe at ballot b. In practice, S will either contain a single message, or else will have a message for each possible value, indicating that all values are safe. In the first case, the leader will immediately send a $2a$ message with the value contained in that single message. (Both logical messages will be sent in the same physical message.) In the latter case, the leader is informing the acceptors that all values are safe. (All those logical messages will, of course, be encoded in a single physical message.)

$$\begin{array}{l}
\textbf{macro } Phase1c(S)\{ \\
\quad \textbf{when } \forall\, v \in S : \exists\, Q \in Quorum : ShowsSafeAt(Q,\ self,\ v)\,\textbf{;} \\
\quad SendSetOfMessages(\{[type \mapsto \text{“1c”},\ bal \mapsto self,\ val \mapsto v] : v \in S\}) \\
\}
\end{array}$$

The ballot *self* leader can perform a $Phase2a(v)$ action, sending a $2a$ message for value v, if it has not already sent a $2a$ message (for this ballot) and it has sent a ballot *self* $1c$ message with *val* field v.

$$\begin{array}{l}
\textbf{macro } Phase2a(v)\{ \\
\quad \textbf{when } \land\, sentMsgs(\text{“2a”},\ self) = \{\} \\
\qquad\qquad \land\, [type \mapsto \text{“1c”},\ bal \mapsto self,\ val \mapsto v] \in msgs\,\textbf{;} \\
\quad SendMessage([type \mapsto \text{“2a”},\ bal \mapsto self,\ val \mapsto v]) \\
\}
\end{array}$$

The $Phase2b(b)$ action is executed by acceptor *self* in response to a ballot-b $2a$ message. Note this action can be executed multiple times by the acceptor, but after the first one, all subsequent executions are stuttering steps that do not change the value of any variable.

$$\begin{array}{l}
\textbf{macro } Phase2b(b)\{ \\
\quad \textbf{when } b \geq maxBal[self]\,\textbf{;} \\
\quad \textbf{with } (m \in sentMsgs(\text{“2a”},\ b))\{ \\
\qquad maxBal[self] \ \ := b\,\textbf{;} \\
\qquad maxVBal[self] := b\,\textbf{;} \\
\qquad maxVVal[self] := m.val\,\textbf{;} \\
\qquad SendMessage([type \mapsto \text{“2b”},\ acc \mapsto self,\ bal \mapsto b,\ val \mapsto m.val]) \\
\quad \}
\end{array}$$

}

An acceptor performs the body of its *while* loop as a single atomic action by nondeterministically choosing a ballot in which its $Phase1b$ or $Phase2b$ action is enabled and executing that enabled action. If no such action is enabled, the acceptor does nothing.

process $(acceptor \in Acceptor)\{$
  acc: **while** (TRUE){
      with $(b \in Ballot)\{$**either** $Phase1b(b)$**or** $Phase2b(b)$
      }
  }
}

The leader of a ballot nondeterministically chooses one of its actions that is enabled (and the argument for which it is enabled) and performs it atomically. It does nothing if none of its actions is enabled.

process $(leader \in Ballot)\{$
  ldr: **while** (TRUE){
      either $Phase1a()$
      or **with** $(S \in \text{SUBSET } Value)\{Phase1c(S)\}$
      or **with** $(v \in Value)\{Phase2a(v)\}$
      }
  }

}

The translator produces the following TLA+ specification of the algorithm. Some blank lines have been deleted.
****

BEGIN TRANSLATION

VARIABLES $maxBal, maxVBal, maxVVal, msgs$

define statement

$sentMsgs(t, b) \triangleq \{m \in msgs : (m.type = t) \land (m.bal = b)\}$

$$
\begin{aligned}
&ShowsSafeAt(Q, b, v) \triangleq \\
&\quad \text{LET } Q1b \triangleq \{m \in sentMsgs(\text{“1b”}, b) : m.acc \in Q\} \\
&\quad \text{IN } \land \forall a \in Q : \exists m \in Q1b : m.acc = a \\
&\qquad\quad \land \lor \forall m \in Q1b : m.mbal = -1 \\
&\qquad\qquad \lor \exists m1c \in msgs : \\
&\qquad\qquad\quad \land m1c = [type \mapsto \text{“1c”}, bal \mapsto m1c.bal, val \mapsto v] \\
&\qquad\qquad\quad \land \forall m \in Q1b : \land m1c.bal \geq m.mbal \\
&\qquad\qquad\qquad\qquad\qquad \land (m1c.bal = m.mbal) \Rightarrow (m.mval = v)
\end{aligned}
$$

$vars \triangleq \langle maxBal, maxVBal, maxVVal, msgs \rangle$

$ProcSet \triangleq (Acceptor) \cup (Ballot)$

$Init \triangleq$ Global variables

$$
\begin{array}{l}
\land\ maxBal = [a \in Acceptor \mapsto\ -1] \\
\land\ maxVBal = [a \in Acceptor \mapsto\ -1] \\
\land\ maxVVal = [a \in Acceptor \mapsto None] \\
\land\ msgs = \{\}
\end{array}
$$

$$
\begin{array}{l}
acceptor(self) \triangleq \exists\, b \in Ballot : \\
\quad \lor\ \land\ (b > maxBal[self]) \land (sentMsgs(\text{“1a”}, b) \neq \{\}) \\
\quad\quad \land\ maxBal' = [maxBal \text{ EXCEPT } ![self] = b] \\
\quad\quad \land\ msgs' = (msgs \cup \{([type \mapsto \text{“1b”}, acc \mapsto self, bal \mapsto b, \\
\quad\quad\quad\quad mbal \mapsto maxVBal[self], mval \mapsto maxVVal[self]])\}) \\
\quad\quad \land\ \text{UNCHANGED } \langle maxVBal, maxVVal \rangle \\
\quad \lor\ \land\ b \geq maxBal[self] \\
\quad\quad \land\ \exists\, m \in sentMsgs(\text{“2a”}, b) : \\
\quad\quad\quad \land\ maxBal' = [maxBal \text{ EXCEPT } ![self] = b] \\
\quad\quad\quad \land\ maxVBal' = [maxVBal \text{ EXCEPT } ![self] = b] \\
\quad\quad\quad \land\ maxVVal' = [maxVVal \text{ EXCEPT } ![self] = m.val] \\
\quad\quad\quad \land\ msgs' = (msgs \cup \{([type \mapsto \text{“2b”}, acc \mapsto self, bal \mapsto b, val \mapsto m.val])\})
\end{array}
$$

$$
\begin{array}{l}
leader(self) \triangleq\ \land\ \lor\ \land\ msgs' = (msgs \cup \{([type \mapsto \text{“1a”}, bal \mapsto self])\}) \\
\quad\quad \lor\ \land\ \exists\, S \in \text{SUBSET } Value : \\
\quad\quad\quad\quad \land\ \forall\, v \in S : \exists\, Q \in Quorum : ShowsSafeAt(Q, self, v) \\
\quad\quad\quad\quad \land\ msgs' = (msgs \cup (\{[type \mapsto \text{“1c”}, bal \mapsto self, val \mapsto v] : v \in S\})) \\
\quad\quad \lor\ \land\ \exists\, v \in Value : \\
\quad\quad\quad\quad \land\ \land\ sentMsgs(\text{“2a”}, self) = \{\} \\
\quad\quad\quad\quad\quad \land\ [type \mapsto \text{“1c”}, bal \mapsto self, val \mapsto v] \in msgs \\
\quad\quad\quad\quad \land\ msgs' = (msgs \cup \{([type \mapsto \text{“2a”}, bal \mapsto self, val \mapsto v])\}) \\
\quad \land\ \text{UNCHANGED } \langle maxBal, maxVBal, maxVVal \rangle
\end{array}
$$

$$
\begin{array}{l}
Next \triangleq (\exists\, self \in Acceptor : acceptor(self)) \\
\quad\quad\quad \lor (\exists\, self \in Ballot : leader(self))
\end{array}
$$

$$
Spec \triangleq Init \land \Box[Next]_{vars}
$$

END TRANSLATION

We now rewrite the next-state relation in a way that makes it easier to use in a proof. We start by defining the formulas representing the individual actions. We then use them to define the formula *TLANext*, which is the next-state relation we would have written had we specified the algorithm directly in TLA+ rather than in *PlusCal*.

$$
\begin{array}{l}
Phase1a(self) \triangleq \\
\quad \land\ msgs' = (msgs \cup \{[type \mapsto \text{“1a”}, bal \mapsto self]\}) \\
\quad \land\ \text{UNCHANGED } \langle maxBal, maxVBal, maxVVal \rangle
\end{array}
$$

$$
\begin{array}{l}
Phase1c(self, S) \triangleq \\
\quad \land\ \forall\, v \in S : \exists\, Q \in Quorum : ShowsSafeAt(Q, self, v) \\
\quad \land\ msgs' = (msgs \cup \{[type \mapsto \text{“1c”}, bal \mapsto self, val \mapsto v] : v \in S\})
\end{array}
$$

$\wedge$ UNCHANGED $\langle maxBal, maxVBal, maxVVal\rangle$

$Phase2a(self, v) \triangleq$
$\wedge$ $sentMsgs(\text{"2a"}, self) = \{\}$
$\wedge$ $[type \mapsto \text{"1c"}, bal \mapsto self, val \mapsto v] \in msgs$
$\wedge$ $msgs' = (msgs \cup \{[type \mapsto \text{"2a"}, bal \mapsto self, val \mapsto v]\})$
$\wedge$ UNCHANGED $\langle maxBal, maxVBal, maxVVal\rangle$

$Phase1b(self, b) \triangleq$
$\wedge$ $b > maxBal[self]$
$\wedge$ $sentMsgs(\text{"1a"}, b) \neq \{\}$
$\wedge$ $maxBal' = [maxBal$ EXCEPT $![self] = b]$
$\wedge$ $msgs' = msgs \cup \{[type \mapsto \text{"1b"}, acc \mapsto self, bal \mapsto b,$
$mbal \mapsto maxVBal[self], mval \mapsto maxVVal[self]]\}$
$\wedge$ UNCHANGED $\langle maxVBal, maxVVal\rangle$

$Phase2b(self, b) \triangleq$
$\wedge$ $b \geq maxBal[self]$
$\wedge$ $\exists m \in sentMsgs(\text{"2a"}, b)$:
$\wedge$ $maxBal' = [maxBal$ EXCEPT $![self] = b]$
$\wedge$ $maxVBal' = [maxVBal$ EXCEPT $![self] = b]$
$\wedge$ $maxVVal' = [maxVVal$ EXCEPT $![self] = m.val]$
$\wedge$ $msgs' = (msgs \cup \{[type \mapsto \text{"2b"}, acc \mapsto self,$
$bal \mapsto b, val \mapsto m.val]\})$

$TLANext \triangleq$
$\vee$ $\exists self \in Acceptor$:
$\exists b \in Ballot$: $\vee$ $Phase1b(self, b)$
$\vee$ $Phase2b(self, b)$
$\vee$ $\exists self \in Ballot$:
$\vee$ $Phase1a(self)$
$\vee$ $\exists S \in$ SUBSET $Value : Phase1c(self, S)$
$\vee$ $\exists v \in Value : Phase2a(self, v)$

The following theorem specifies the relation between the next-state relation *Next* obtained by translating the *PlusCal* code and the next-state relation *TLANext*.

THEOREM $NextDef \triangleq (Next \equiv TLANext)$
$\langle 1\rangle 2.$ ASSUME NEW $self \in Acceptor$
PROVE $acceptor(self) \equiv TLANext!1!(self)$
BY $\langle 1\rangle 2$, $BallotAssump$ DEF $acceptor$, $ProcSet$, $Phase1b$, $Phase2b$
$\langle 1\rangle 3.$ ASSUME NEW $self \in Ballot$
PROVE $leader(self) \equiv TLANext!2!(self)$
BY $\langle 1\rangle 3$, $BallotAssump$, $Zenon$ DEF $leader$, $ProcSet$, $Phase1a$, $Phase1c$, $Phase2a$
$\langle 1\rangle 4.$ QED
BY $\langle 1\rangle 2$, $\langle 1\rangle 3$ DEF $Next$, $TLANext$

The type invariant.

$$TypeOK \triangleq \land maxBal \in [Acceptor \rightarrow Ballot \cup \{-1\}]$$
$$\land maxVBal \in [Acceptor \rightarrow Ballot \cup \{-1\}]$$
$$\land maxVVal \in [Acceptor \rightarrow Value \cup \{None\}]$$
$$\land msgs \subseteq Message$$

Here is the definition of the state-function *chosen* that implements the state-function of the same name in the voting algorithm.

$$chosen \triangleq \{v \in Value : \exists Q \in Quorum, b \in Ballot : \forall a \in Q : \exists m \in msgs : \land m.type = \text{"2b"} \land m.acc = a \land m.bal = b \land m.val = v\}$$

We now define the refinement mapping under which this algorithm implements the specification in module Voting.

As we observed, votes are registered by sending phase $2b$ messages. So the array *votes* describing the votes cast by the acceptors is defined as follows.

$$votes \triangleq [a \in Acceptor \mapsto \{\langle m.bal, m.val \rangle : m \in \{mm \in msgs : \land mm.type = \text{"2b"} \land mm.acc = a\}\}]$$

We now instantiate module Voting, substituting:

- The constants *Value*, *Acceptor*, and *Quorum* declared in this module for the corresponding constants of that module Voting.
- The variable *maxBal* and the defined state function *votes* for the correspondingly-named variables of module Voting.

$$V \triangleq \text{INSTANCE } VoteProof$$

We now define *PInv* to be what I believe to be an inductive invariant and assert the theorems for proving that this algorithm implements the voting algorithm under the refinement mapping specified by the INSTANCE statement. Whether *PInv* really is an inductive invariant will be determined only by a rigorous proof.

$$PAccInv \triangleq \forall a \in Acceptor :$$
$$\land maxBal[a] \geq maxVBal[a]$$
$$\land \forall b \in (maxVBal[a] + 1) \,..\, (maxBal[a] - 1) : V!DidNotVoteIn(a, b)$$
$$\land (maxVBal[a] \neq -1) \Rightarrow V!VotedFor(a, maxVBal[a], maxVVal[a])$$

$$P1bInv \triangleq \forall m \in msgs :$$
$$(m.type = \text{"1b"}) \Rightarrow$$
$$\land (maxBal[m.acc] \geq m.bal) \land (m.bal > m.mbal)$$
$$\land \forall b \in (m.mbal + 1) \,..\, (m.bal - 1) : V!DidNotVoteIn(m.acc, b)$$

$$P1cInv \triangleq \forall m \in msgs : (m.type = \text{"1c"}) \Rightarrow V!SafeAt(m.bal, m.val)$$

$P2aInv \triangleq \forall m \in msgs :$
$$(m.type = \text{“2a”}) \Rightarrow \exists m1c \in msgs : \land m1c.type = \text{“1c”} \land m1c.bal = m.bal \land m1c.val = m.val$$

The following theorem is interesting in its own right. It essentially asserts the correctness of the definition of *ShowsSafeAt*.

THEOREM $PT1 \triangleq TypeOK \land P1bInv \land P1cInv \Rightarrow \forall Q \in Quorum, b \in Ballot, v \in Value : ShowsSafeAt(Q, b, v) \Rightarrow V!SafeAt(b, v)$

$PInv \triangleq TypeOK \land PAccInv \land P1bInv \land P1cInv \land P2aInv$

THEOREM $Invariance \triangleq Spec \Rightarrow \Box PInv$

THEOREM $Implementation \triangleq Spec \Rightarrow V!Spec$

The following result shows that our definition of *chosen* is the correct one, because it implements the state-function *chosen* of the voting algorithm.

THEOREM $Spec \Rightarrow \Box(chosen = V!chosen)$

The four theorems above have been checked by *TLC* for a model with 3 acceptors, 2 values, and 3 ballot numbers. Theorem $PT1$ was checked as an invariant, therefore checking only that it is true for all reachable states. This model is large enough that it would most likely have revealed any “coding” errors in the algorithm. We believe that the algorithm is well-enough understood that it is unlikely to contain any fundamental errors.

* Modification History
* Last modified *Fri* May 22 09:20:18 *CEST* 2020 by *merz*
* Last modified *Fri* *Jul* 15 11:31:15 *PDT* 2011 by lamport

---

```
(*************************************************************************)
(*                          Liveness                                     *)
(*                                                                       *)
(* The liveness property satisfied by PCon (and classic Paxos) is:      *)
(*                                                                       *)
(* If there is some ballot b and quorum Q such that                     *)
(*                                                                       *)
(* 1. No phase 1a messages (a) have been or (b) ever will be sent for any *)
(*    ballot number greater than b.                                      *)
(*                                                                       *)
(* 2. The ballot b leader eventually sends a phase 1a message for ballot *)
(*    b.                                                                 *)
(*                                                                       *)
(* 3. Each acceptor in Q eventually responds to ballot b messages sent  *)
(*    by the ballot b leader–which implies that it eventually receives  *)
(*    those messages.                                                    *)
(*                                                                       *)
(* 4. The ballot b leader eventually executes its Phase2a action for    *)
(*    ballot b if it can.                                                *)
(*                                                                       *)
```

```
(* then some value is eventually chosen.                         *)
(*                                                               *)
(* Note that Phase2a(b) is enabled if msgs contains a ballot b phase 1b *)
(* message from every acceptor in Q. Hence, 4 implies that if the leader *)
(* eventually receives those messages, then it must perform its Phase2a(b) *)
(* action. (It might perform that action before it receives those      *)
(* messages if it has received phase 1b messages from all the acceptors in *)
(* a different quorum.)                                          *)
(***************************************************************************)
THEOREM Liveness ≜
 Spec ⇒ ∀ b ∈ Ballot, Q ∈ Quorum :
        ( ( ∧ (************************************************************)
              (* Assumption 1a.                                         *)
              (************************************************************) ∀ m ∈ msgs :
              (m.type = "1a") ⇒ (m.bal < b)
            ∧ (************************************************************)
              (* Assumption 1b.                                         *)
              (************************************************************) ∀ c ∈ Ballot :
              (c > b) ⇒ □[¬Phase1a(c)]_vars
            ∧ (************************************************************)
              (* Assumption 2.                                          *)
              (************************************************************)
              WF_vars(Phase1a(b))
            ∧ (************************************************************)
              (* Assumption 4.                                          *)
              (************************************************************)
              WF_vars(∃ v ∈ Value : Phase2a(b, v))
            ∧ (************************************************************)
              (* Assumption 3.                                          *)
              (************************************************************)
              ∀ a ∈ Q : ∧ WF_vars(Phase1b(a, b))
                        ∧ WF_vars(Phase2b(a, b))
          ) ⇝ (chosen ≠ {}) )
```

```
\* The following is used to check theorem Liveness

CONSTANTS bb, QQ

CSpec ≜ ∧ Init
        ∧ □[ ∧ Next
             ∧ ∀ c ∈ Ballot : (c > bb) ⇒ ¬Phase1a(c)]_vars
        ∧ WF_vars(Phase1a(bb))
        ∧ WF_vars(∃ v ∈ Value : Phase2a(bb, v))
        ∧ ∀ a ∈ QQ : ∧ WF_vars(Phase1bForBallot(a, bb))
                     ∧ WF_vars(Phase2bForBallot(a, bb))

CLiveness ≜ (∀ m ∈ msgs : (m.type = "1a") ⇒ (m.bal < bb)) ⇝ (chosen ≠ {})
```