The Paxos Consensus Algorithm

Remember that the voting algorithm is based on two rules:

Remember that the voting algorithm is based on two rules:

1. Don't allow different acceptors to vote for different values in the same ballot.

Remember that the voting algorithm is based on two rules:

- 1. Don't allow different acceptors to vote for different values in the same ballot.
- 2. Allow an acceptor to vote for value v in ballot b only if v is safe at b.

- 1. Don't allow different acceptors to vote for different values in the same ballot.
- 2. Allow an acceptor to vote for value v in ballot b only if
 - v is safe at b.

- 1. Don't allow different acceptors to vote for different values in the same ballot.
- 2. Allow an acceptor to vote for value v in ballot b only if v is safe at b.

- 1. Don't allow different acceptors to vote for different values in the same ballot.
- 2. Allow an acceptor to vote for value v in ballot b only if v is safe at b.

- 1. Don't allow different acceptors to vote for different values in the same ballot.
- 2. Allow an acceptor to vote for value v in ballot b only if v is safe at b.

We don't care who the leaders are.

- 1. Don't allow different acceptors to vote for different values in the same ballot.
- 2. Allow an acceptor to vote for value v in ballot b only if v is safe at b.

We don't care who the leaders are.

The same process may be the leader for many different ballots.

- 1. Don't allow different acceptors to vote for different values in the same ballot.
- 2. Allow an acceptor to vote for value v in ballot b only if v is safe at b.

The leader of a ballot tells acceptors what value to vote for in that ballot.

- 1. Don't allow different acceptors to vote for different values in the same ballot.
- 2. Allow an acceptor to vote for value v in ballot b only if
 - v is safe at b.

- 1. Don't allow different acceptors to vote for different values in the same ballot.
- 2. Allow an acceptor to vote for value v in ballot b only if v is safe at b.

- 1. Don't allow different acceptors to vote for different values in the same ballot.
- 2. Allow an acceptor to vote for value v in ballot b only if v is safe at b.

If it receives responses from a quorum of acceptors,

- 1. Don't allow different acceptors to vote for different values in the same ballot.
- 2. Allow an acceptor to vote for value v in ballot b only if v is safe at b.

If it receives responses from a quorum of acceptors, it will be able to choose a safe value and tell them to vote for it.

- 1. Don't allow different acceptors to vote for different values in the same ballot.
- 2. Allow an acceptor to vote for value v in ballot b only if v is safe at b.

If it receives responses from a quorum of acceptors, it will be able to choose a safe value and tell them to vote for it.

Let's look at the spec.

——— MODULE Paxos ————

— MODULE Paxos —

EXTENDS Integers

— MODULE Paxos —

EXTENDS Integers

CONSTANTS Value, Acceptor, Quorum

_____ MODULE Paxos _____

EXTENDS Integers

CONSTANTS Value, Acceptor, Quorum

ASSUME $\land \forall Q \in Quorum : Q \subseteq Acceptor$ $\land \forall Q1, Q2 \in Quorum : Q1 \cap Q2 \neq \{\}$

EXTENDS Integers MODULE Paxos —

CONSTANTS Value, Acceptor, Quorum

ASSUME $\land \forall Q \in Quorum : Q \subseteq Acceptor$ $\land \forall Q1, Q2 \in Quorum : Q1 \cap Q2 \neq \{\}$

 $Ballot \stackrel{\Delta}{=} Nat$

EXTENDS Integers MODULE Paxos —

CONSTANTS Value, Acceptor, Quorum

ASSUME $\land \forall Q \in Quorum : Q \subseteq Acceptor$ $\land \forall Q1, Q2 \in Quorum : Q1 \cap Q2 \neq \{\}$

 $Ballot \stackrel{\Delta}{=} Nat$

EXTENDS Integers MODULE Paxos —

CONSTANTS Value, Acceptor, Quorum

ASSUME $\land \forall Q \in Quorum : Q \subseteq Acceptor$ $\land \forall Q1, Q2 \in Quorum : Q1 \cap Q2 \neq \{\}$

 $Ballot \stackrel{\Delta}{=} Nat$

None $\stackrel{\Delta}{=}$ CHOOSE $v: v \notin Ballot$

None
$$\stackrel{\Delta}{=}$$
 CHOOSE $v: v \notin Ballot$

This defines None to be some value that is not an element of the set Ballot.

None $\stackrel{\Delta}{=}$ CHOOSE $v: v \notin Ballot$

This defines *None* to be some value that is not an element of the set *Ballot*.

We don't know (or care) what value.

This defines *None* to be some value that is not an element of the set *Ballot*.

We don't know (or care) what value.

This is a mathematical expression,

This defines *None* to be some value that is not an element of the set *Ballot*.

We don't know (or care) what value.

This is a mathematical expression, so evaluating it for the same value of *Ballot* always yields the same result.

This defines *None* to be some value that is not an element of the set *Ballot*.

We don't know (or care) what value.

This is a mathematical expression, so evaluating it for the same value of *Ballot* always yields the same result.

And *Ballot* is a constant, so its value doesn't change,

This defines *None* to be some value that is not an element of the set *Ballot*.

We don't know (or care) what value.

This is a mathematical expression, so evaluating it for the same value of *Ballot* always yields the same result.

And *Ballot* is a constant, so its value doesn't change, and therefore, neither does the value of *None*.

$Message \stackrel{\Delta}{=}$

$Message \stackrel{\Delta}{=}$

The leaders and acceptors will communicate with messages.



The leaders and acceptors will communicate with messages.

Message is defined to be the set of all messages that can ever be sent.

$Message \stackrel{\Delta}{=}$

The leaders and acceptors will communicate with messages.

Message is defined to be the set of all messages that can ever be sent.

A message is represented by a record.

Records

Records

A record is similar to a struct in C,
A record is similar to a struct in C, or an object in an OO language,

A record is similar to a *struct* in C, or an *object* in an OO language, with fields but no concept of methods.

It is defined mathematically to be a function whose domain is a finite set of strings.

It is defined mathematically to be a function whose domain is a finite set of strings.

For example this record

It is defined mathematically to be a function whose domain is a finite set of strings.

For example this record

 $[type \mapsto "2a",$

with type field equal to "2a"

It is defined mathematically to be a function whose domain is a finite set of strings.

For example this record

 $[type \mapsto "2a", bal \mapsto 42,]$

with type field equal to "2a", bal field equal to 42

It is defined mathematically to be a function whose domain is a finite set of strings.

For example this record

 $[type \mapsto "2a", bal \mapsto 42, val \mapsto \sqrt{7}]$

with type field equal to "2*a*", *bal* field equal to 42, and *val* field equal to $\sqrt{7}$

It is defined mathematically to be a function whose domain is a finite set of strings.

For example this record

 $[type \mapsto "2a", bal \mapsto 42, val \mapsto \sqrt{7}]$

It is defined mathematically to be a function whose domain is a finite set of strings.

For example this record

 $[type \mapsto "2a", bal \mapsto 42, val \mapsto \sqrt{7}]$

is the function

 $[\,x \in \{"type","bal","val"\} \mapsto$

It is defined mathematically to be a function whose domain is a finite set of strings.

For example this record

 $[type \mapsto "2a", bal \mapsto 42, val \mapsto \sqrt{7}]$

is the function

 $[\,x \in \{"type","bal","val"\} \mapsto$

It is defined mathematically to be a function whose domain is a finite set of strings.

For example this record

 $[type \mapsto "2a", bal \mapsto 42, val \mapsto \sqrt{7}]$

is the function

 $[\,x \in \{"type","bal","val"\} \mapsto$

It is defined mathematically to be a function whose domain is a finite set of strings.

For example this record

 $[type \mapsto "2a", bal \mapsto 42, val \mapsto \sqrt{7}]$

is the function

 $[x \in \{"type", "bal", "val"\} \mapsto \\ \mathsf{IF} \ x = "type" \ \mathsf{THEN} \ "2a" \\ \end{cases}$

It is defined mathematically to be a function whose domain is a finite set of strings.

For example this record

 $[type \mapsto "2a", bal \mapsto 42, val \mapsto \sqrt{7}]$

$$[x \in \{"type", "bal", "val"\} \mapsto \\ IF \ x = "type" \ THEN \ "2a" \\ ELSE \ IF \ x = "bal" \ THEN \ 42$$

It is defined mathematically to be a function whose domain is a finite set of strings.

For example this record

$$[type \mapsto "2a", bal \mapsto 42, val \mapsto \sqrt{7}]$$

$$[x \in \{"type", "bal", "val"\} \mapsto \\ \text{IF } x = "type" \text{ THEN "2a"} \\ \text{ELSE IF } x = "bal" \text{ THEN 42} \\ \text{ELSE } \sqrt{7}]$$

It is defined mathematically to be a function whose domain is a finite set of strings.

For example this record

 $[type \mapsto "2a", bal \mapsto 42, val \mapsto \sqrt{7}]$

$$[x \in \{"type", "bal", "val"\} \mapsto \\ \text{IF } x = "type" \text{ THEN "2a"} \\ \text{ELSE IF } x = "bal" \text{ THEN 42} \\ \text{ELSE } \sqrt{7}]$$

 $[type \mapsto "2a", bal \mapsto 42, val \mapsto \sqrt{7}]$

 $[\textit{type} \mapsto \texttt{``2a"}, \textit{ bal} \mapsto \texttt{42}, \textit{ val} \mapsto \sqrt{7}][\texttt{``bal"}] = \texttt{42}$

 $[type \mapsto "2a", bal \mapsto 42, val \mapsto \sqrt{7}]["bal"]$

$[type \mapsto "2a", bal \mapsto 42, val \mapsto \sqrt{7}]$ ["bal"]

We abbreviation ["bal"]

[$type \mapsto$ "2a", $bal \mapsto$ 42, $val \mapsto \sqrt{7}$]. bal

We abbreviation ["bal"] as .bal

$[type \mapsto "2a", bal \mapsto 42, val \mapsto \sqrt{7}].bal = 42$

We abbreviation ["bal"] as .bal

$Message \stackrel{\Delta}{=}$

Defined to be a set of records.

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

Defined to be a set of records.



Defined to be a set of records.

The union of four sets.







The set of all records R having three fields type, bal, and val



The set of all records R having three fields type, bal, and val with: R.type in the set {"2a"}



The set of all records R having three fields type, bal, and val with: R.type in the set {"2a"}

R.bal in the set *Ballot*



The set of all records R having three fields type, bal, and val with: R.type in the set {"2a"}

- *R.bal* in the set *Ballot*
- *R.val* in the set *Value*



The set of all records R having three fields type, bal, and val with:

R.type in the set {"2*a*"}

- *R.bal* in the set *Ballot*
- *R.val* in the set *Value*

$$\begin{array}{rcl}Message \stackrel{\Delta}{=} & & & \\ & & \begin{bmatrix} & & \\ & &$$

The set of all records R having three fields type, bal, and val with:

R.type = "2a"

- *R.bal* in the set *Ballot*
- *R.val* in the set *Value*



VARIABLES

VARIABLES maxBal

Implements variable maxBal of the Voting spec.
VARIABLES maxBal

Implements variable *maxBal* of the *Voting* spec.

maxBal[a] is the number of the highest-numbered ballot in which acceptor a has participated.

VARIABLES maxBal

Implements variable *maxBal* of the *Voting* spec.

maxBal[a] is the number of the highest-numbered ballot in which acceptor a has participated.

Or -1 if *a* hasn't participated in any ballots.

VARIABLES maxBal, maxVBal

VARIABLES maxBal, maxVBal

maxVBal[a] is the number of the highest-numbered ballot in which acceptor a has voted.

VARIABLES maxBal, maxVBal

maxVBal[a] is the number of the highest-numbered ballot in which acceptor a has voted.

Or -1 if *a* hasn't voted in any ballots.

VARIABLES maxBal, maxVBal, maxVal

VARIABLES maxBal, maxVBal, maxVal

maxVal[a] is the value *a* voted for in ballot maxVBal[a].

VARIABLES maxBal, maxVBal, maxVal

maxVal[a] is the value *a* voted for in ballot maxVBal[a].

Or *None* if a has never voted.

VARIABLES maxBal, maxVBal, maxVal, msgs

VARIABLES maxBal, maxVBal, maxVal, msgs

VARIABLES maxBal, maxVBal, maxVal, msgs

msgs

msgs

A message is sent by adding it to msgs.

A message is sent by adding it to *msgs*.

Messages are never removed from msgs.

A message is sent by adding it to msgs.

Messages are never removed from msgs.

This is the simplest, most abstract way I know to describe message passing.

msgs

It can be used to model the Paxos algorithm because:

- Paxos allows messages to be received in any order.

- Paxos allows messages to be received in any order.
- Since we're not specifying liveness, there's no need to explicitly describe message loss.

- Paxos allows messages to be received in any order.
- Since we're not specifying liveness, there's no need to explicitly describe message loss.
 - A spec of safety doesn't require any message to be received,

- Paxos allows messages to be received in any order.
- Since we're not specifying liveness, there's no need to explicitly describe message loss.

A spec of safety doesn't require any message to be received, and there's no difference between a lost message and one that's never received.

msgs

This is a natural, obvious way to describe message passing

This is a natural, obvious way to describe message passing if you think mathematically.

This is a natural, obvious way to describe message passing if you think mathematically.

It doesn't occur to most people because they think in terms of programming languages, not math.

$vars \stackrel{\Delta}{=} \langle maxBal, maxVBal, maxVal, msgs \rangle$

 $vars \stackrel{\Delta}{=} \langle maxBal, maxVBal, maxVal, msgs \rangle$

It's easier to write $[Next]_{vars}$

$vars \stackrel{\Delta}{=} \langle maxBal, maxVBal, maxVal, msgs \rangle$

It's easier to write [Next]_{vars}

than $[Next]_{\langle maxBal, maxVBal, maxVal, msgs \rangle}$.

$TypeOK \stackrel{\Delta}{=}$

11

$TypeOK \stackrel{\Delta}{=} \land maxBal \quad \in [Acceptor \rightarrow Ballot \cup \{-1\}]$

maxBal is a function from acceptors to ballot numbers or -1.

$\begin{array}{rl} TypeOK \ \triangleq \ \land maxBal \ \in [Acceptor \rightarrow Ballot \cup \{-1\}] \\ \land maxVBal \in [Acceptor \rightarrow Ballot \cup \{-1\}] \end{array}$

maxBal is a function from acceptors to ballot numbers or -1. maxVBal is a function from acceptors to ballot numbers or -1.

maxBal is a function from acceptors to ballot numbers or -1. maxVBal is a function from acceptors to ballot numbers or -1.

maxVal is a function from acceptors to values or *None*.

maxBal is a function from acceptors to ballot numbers or -1. maxVBal is a function from acceptors to ballot numbers or -1. maxVal is a function from acceptors to values or *None*. msgs is a subset of the set of all possible messages.

$\mathit{Init} \; \stackrel{\scriptscriptstyle \Delta}{=} \;$

11

$$Init \stackrel{\Delta}{=} \land maxBal = [a \in Acceptor \mapsto -1]$$

For all a : maxBal[a] = -1

$$\begin{array}{rcl} Init \ \triangleq \ \land maxBal &= [a \in Acceptor \mapsto \ -1] \\ \land maxVBal &= [a \in Acceptor \mapsto \ -1] \end{array}$$

For all a : maxBal[a] = -1maxVBal[a] = -1
$$\begin{array}{rcl} Init & \triangleq & \wedge maxBal & = [a \in Acceptor \mapsto -1] \\ & \wedge maxVBal = [a \in Acceptor \mapsto -1] \\ & \wedge maxVal & = [a \in Acceptor \mapsto None] \end{array}$$

For all a : maxBal[a] = -1 maxVBal[a] = -1maxVal[a] = None

$$\begin{array}{rll} Init \ &\triangleq \ \land maxBal \ = [a \in Acceptor \mapsto -1] \\ \land maxVBal = [a \in Acceptor \mapsto -1] \\ \land maxVal \ = [a \in Acceptor \mapsto -1] \\ \land masSal \ = [a \in Acceptor \mapsto None] \\ \land msgs = \{\} \end{array}$$

For all
$$a : maxBal[a] = -1$$

 $maxVBal[a] = -1$
 $maxVal[a] = None$

 $msgs = \{ \}$

14

Phase 1a(b)

Phase 1a(b)

The ballot b leader sends a message asking acceptors to participate in that ballot.

Phase 1a(b)

Phase 1b(a)

Phase 1a(b)

Phase 1b(a)

If acceptor a has not agreed to participate in a ballot numbered $\geq b$, then it agrees and sends the leader maxBal[a], maxVBal[a], and maxVal[a].

Phase 1a(b)

Phase 1b(a)

Phase 2a(b, v)

Phase 1a(b)

Phase1b(a)

Phase 2a(b, v)

If the ballot b leader receives those messages from a quorum, it chooses a value v safe at b and sends a message asking acceptors to vote for v in ballot b.

Phase 1a(b)

Phase 1b(a)

Phase 2a(b, v)

Phase 2b(a)

Phase 1a(b)

Phase 1b(a)

```
Phase 2a(b, v)
```

```
Phase 2b(a)
```

If acceptor a has not agreed to participate in a ballot numbered > b, then it votes for v in ballot b.

Phase 1a(b)

Phase 1b(a)

Phase 2a(b, v)

Phase 2b(a)

If acceptor a has not agreed to participate in a ballot numbered > b, then it votes for v in ballot b.

It does this by sending a message.

Phase 1a(b)

Phase 1b(a)

Phase 2a(b, v)

Phase 2b(a)

If acceptor a has not agreed to participate in a ballot numbered > b, then it votes for v in ballot b.

It does this by sending a message.

We don't care who receives that message.

$Send(m) \stackrel{\Delta}{=}$



Describes the sending of a message m.

 $Send(m) \stackrel{\Delta}{=} msgs' = msgs \cup \{m\}$

Describes the sending of a message m.

$$Phase1a(b) \stackrel{\Delta}{=}$$

 $Phase1a(b) \stackrel{\Delta}{=}$

The ballot b leader sends a message asking acceptors to participate in that ballot.

$$Phase1a(b) \stackrel{\Delta}{=}$$

$Phase1a(b) \stackrel{\Delta}{=} \land Send([type \mapsto "1a", bal \mapsto b])$

$$Phase1a(b) \stackrel{\Delta}{=} \land Send([type \mapsto "1a", bal \mapsto b])$$

Sends a type 1*a* message containing the ballot number.

Leaves all variables except *msgs* unchanged.

$$Phase1b(a) \stackrel{\Delta}{=}$$

$$Phase1b(a) \stackrel{\Delta}{=}$$

If acceptor a has not agreed to participate in a ballot numbered $\geq b$,

$$Phase1b(a) \stackrel{\Delta}{=}$$

If acceptor a has not agreed to participate in a ballot numbered $\geq b$, then it agrees and sends the leader maxBal[a], maxVBal[a], and maxVal[a].

$$Phase1b(a) \stackrel{\Delta}{=}$$

 $Phase1b(a) \stackrel{\Delta}{=}$

 $Phase1b(a) \stackrel{\Delta}{=}$

$$Phase1b(a) \triangleq \\ \land \exists m \in msgs : \\ \land m.type = "1a"$$

 $\begin{array}{ll} Phase1b(a) &\triangleq \\ & \land \exists \ m \in msgs: \\ & \land \ m.type = \ ``1a'' \\ & \land \ m.bal > maxBal[a] \end{array}$

$\begin{array}{ll} Phase1b(a) &\triangleq \\ & \land \exists \ m \in msgs: \\ & \land \ m.type = \ ``1a'' \\ & \land \ m.bal > maxBal[a] \end{array}$



 $\begin{array}{ll} Phase1b(a) &\triangleq \\ & \land \exists \ m \in msgs: \\ & \land \ m.type = \ ``1a'' \\ & \land \ m.bal > maxBal[a] \end{array}$

$$\begin{array}{l} Phase1b(a) \triangleq \\ \land \exists m \in msgs: \\ \land m.type = ``1a'' \\ \land m.bal > maxBal[a] \\ \land maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \end{array}$$

$$\begin{array}{l} Phase1b(a) \triangleq \\ \land \exists m \in msgs: \\ \land m.type = ``1a'' \\ \land m.bal > maxBal[a] \\ \land maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \end{array}$$
$$\begin{array}{l} Phase1b(a) \triangleq \\ \land \exists m \in msgs: \\ \land m.type = ``1a'' \\ \land m.bal > maxBal[a] \\ \land maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \\ \end{array}$$

$$\begin{array}{l} \textbf{Setting } maxBal[a] \ \textbf{to } m.bal \ . \end{array}$$

$$\begin{array}{l} Phase1b(a) \triangleq \\ \land \exists m \in msgs: \\ \land m.type = ``1a'' \\ \land m.bal > maxBal[a] \\ \land maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \end{array}$$

$$\begin{array}{l} Phase1b(a) \triangleq \\ \land \exists m \in msgs: \\ \land m.type = ``1a'' \\ \land m.bal > maxBal[a] \\ \land maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \\ \land Send([type \mapsto ``1b'', \ acc \mapsto a, \ bal \mapsto m.bal, \\ mbal \mapsto maxVBal[a], \ mval \mapsto maxVal[a]]) \end{array}$$

$$\begin{array}{ll} Phase1b(a) &\triangleq \\ & \land \exists m \in msgs: \\ & \land m.type = ``1a'' \\ & \land m.bal > maxBal[a] \\ & \land maxBal' = [maxBal \ \texttt{EXCEPT} \ ![a] = m.bal] \\ & \land Send([type \mapsto ``1b'', \ acc \mapsto a, \ bal \mapsto m.bal], \\ & mbal \mapsto maxVBal[a], \ mval \mapsto maxVal[a])) \end{array}$$

$$\begin{array}{ll} Phase1b(a) &\triangleq \\ \land \exists \ m \in msgs: \\ \land \ m.type = ``1a'' \\ \land \ m.bal > maxBal[a] \\ \land \ maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \\ \land \ Send([type \mapsto ``1b'', \ acc \mapsto a, \ bal \mapsto \boxed{m.bal}, \\ mbal \mapsto maxVBal[a], \ mval \mapsto maxVal[a]]) \end{array}$$

$$\begin{array}{l} Phase1b(a) \triangleq \\ \land \exists \ m \in msgs: \\ \land \ m.type = ``1a'' \\ \land \ m.bal > maxBal[a] \\ \land \ maxBal' = [maxBal \ \text{EXCEPT } ![a] = m.bal] \\ \land \ Send([type \mapsto ``1b'', \ acc \mapsto a, \ bal \mapsto m.bal, \\ mbal \mapsto maxVBal[a], \ mval \mapsto maxVal[a]]) \end{array}$$

$$\begin{array}{l} Phase1b(a) \triangleq \\ \land \exists m \in msgs: \\ \land m.type = ``1a'' \\ \land m.bal > maxBal[a] \\ \land maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \\ \land Send([type \mapsto ``1b'', \ acc \mapsto a, \ bal \mapsto m.bal, \\ mbal \mapsto maxVBal[a], \ mval \mapsto maxVal[a]]) \\ \land \text{ UNCHANGED } \langle maxVBal, \ maxVal \rangle \end{array}$$

$$\begin{array}{ll} Phase1b(a) &\triangleq \\ & \land \exists m \in msgs: \\ & \land m.type = ``1a'' \\ & \land m.bal > maxBal[a] \\ & \land maxBal' = [maxBal \ \texttt{EXCEPT} \ ![a] = m.bal] \\ & \land Send([type \mapsto ``1b'', \ acc \mapsto a, \ bal \mapsto m.bal, \\ & mbal \mapsto maxVBal[a], \ mval \mapsto maxVal[a]]) \\ & \land \texttt{UNCHANGED} \ \langle maxVBal, \ maxVal \rangle \end{array}$$

And it leaves the other variables unchanged.

$$\begin{array}{l} Phase1b(a) \triangleq \\ \land \exists m \in msgs: \\ \land m.type = ``1a'' \\ \land m.bal > maxBal[a] \\ \land maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \\ \land Send([type \mapsto ``1b'', \ acc \mapsto a, \ bal \mapsto m.bal, \\ mbal \mapsto maxVBal[a], \ mval \mapsto maxVal[a]]) \\ \land \text{ UNCHANGED } \langle maxVBal, \ maxVal \rangle \end{array}$$

$$Phase2a(b, v) \triangleq$$

$$Phase2a(b, v) \stackrel{\Delta}{=}$$

If the ballot b leader receives those (1 b) messages from a quorum, it chooses a value v safe at b and sends a message asking acceptors to vote for v in ballot b.

$$Phase2a(b, v) \triangleq$$

$$Phase2a(b, v) \stackrel{\Delta}{=}$$

 $\begin{array}{l} Phase2a(b, v) \triangleq \\ \land \neg \exists \ m \in \ msgs: m.type = ``2a'' \land m.bal = b \end{array}$

 $\begin{array}{l} Phase 2a(b, v) \stackrel{\Delta}{=} \\ \land \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \end{array}$

No 2a message for ballot b has already been sent.

$\begin{array}{l} Phase2a(b, v) \triangleq \\ \land \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \\ \land \exists \ Q \in Quorum : \end{array}$

$\begin{array}{l} Phase2a(b, v) \triangleq \\ \land \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \\ \land \exists \ Q \in Quorum : \\ LET \ Q1b \triangleq \end{array}$

IN

 $\begin{array}{l} Phase2a(b, v) \triangleq \\ \land \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \\ \land \exists \ Q \in Quorum : \\ LET \ Q1b \triangleq \\ \\ Locally defines \ Q1b . \end{array}$

IN

IN

$$\begin{array}{l} Phase2a(b, v) \triangleq \\ \land \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \\ \land \exists \ Q \in Quorum : \\ \text{LET } Q1b \triangleq \{m \in msgs : \land m.type = ``1b'' \\ \land m.acc \in Q \\ \land m.bal = b\} \\ Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\} \\ \text{IN} \end{array}$$

$$\begin{array}{l} Phase2a(b, v) \triangleq \\ \land \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \\ \land \exists \ Q \in Quorum : \\ \text{LET } Q1b \triangleq \{m \in msgs : \land m.type = ``1b'' \\ \land m.acc \in Q \\ \land m.bal = b\} \\ Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\} \\ \text{IN } \land \forall \ a \in Q : \exists \ m \in Q1b : m.acc = a \end{array}$$

$$\begin{array}{l} Phase2a(b, v) \triangleq \\ \land \neg \exists \ m \in msgs: m.type = ``2a'' \land m.bal = b \\ \land \exists \ Q \in Quorum: \\ \text{LET} \ Q1b \triangleq \{m \in msgs: \land m.type = ``1b'' \\ \land m.acc \in Q \\ \land m.bal = b\} \\ Q1bv \triangleq \{m \in Q1b: m.mbal \ge 0\} \\ \text{IN} \ \land \forall \ a \in Q: \exists \ m \in Q1b: m.acc = a \\ \text{There is a message in } Q1b \text{ from every} \\ acceptor \text{ in } Q. \end{array}$$

$$\begin{array}{l} Phase2a(b, v) \triangleq \\ \land \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \\ \land \exists \ Q \in Quorum : \\ \text{LET } Q1b \triangleq \{m \in msgs : \land m.type = ``1b'' \\ \land m.acc \in Q \\ \land m.bal = b\} \\ Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\} \\ \text{IN } \land \forall \ a \in Q : \exists \ m \in Q1b : m.acc = a \\ \land \lor Q1bv = \{\} \end{array}$$

$$\begin{array}{l} Phase2a(b, v) \triangleq \\ \land \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \\ \land \exists \ Q \in Quorum : \\ \text{LET } Q1b \triangleq \{m \in msgs : \land m.type = ``1b'' \\ \land m.acc \in Q \\ \land m.bal = b\} \\ Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\} \\ \text{IN } \land \forall \ a \in Q : \exists \ m \in Q1b : m.acc = a \\ \land \lor Q1bv = \{\} \\ \text{Either no acceptor in } Q \text{ reported} \\ \text{that it had yoted.} \end{array}$$

$$\begin{array}{l} Phase2a(b, v) \triangleq \\ \land \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \\ \land \exists \ Q \in Quorum : \\ \text{ LET } Q1b \triangleq \{m \in msgs : \land m.type = ``1b'' \\ \land m.acc \in Q \\ \land m.bal = b\} \\ Q1bv \triangleq \{m \in Q1b : m.mbal \geq 0\} \\ \text{ IN } \land \forall \ a \in Q : \exists \ m \in Q1b : m.acc = a \\ \land \lor Q1bv = \{\} \\ \lor \exists \ m \in Q1bv : \\ \land m.mval = v \\ \land \forall \ mm \in Q1bv : m.mbal \geq mm.mbal \\ \text{ or the message in } Q1bv \ reporting a \\ \text{ vote in the highest-numbered ballot} \\ \text{ reports a vote for } v . \end{array}$$

$$\begin{array}{l} Phase2a(b, v) \triangleq \\ & \wedge \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \\ & \land \exists \ Q \in Quorum : \\ & \text{LET} \ Q1b \triangleq \{m \in msgs : \land m.type = ``1b'' \\ & \land m.acc \in Q \\ & \land m.bal = b\} \\ & Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\} \\ & \text{IN} \quad \land \forall \ a \in Q : \exists \ m \in Q1b : m.acc = a \\ & \land \lor Q1bv = \{\} \\ & \lor \exists \ m \in Q1bv : \\ & \land m.mval = v \\ & \land \forall \ mm \in Q1bv : m.mbal \ge mm.mbal \end{array}$$

This condition implies that v is safe at b

$$\begin{array}{l} Phase2a(b, v) \triangleq \\ \land \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \\ \land \exists \ Q \in Quorum : \\ \\ \text{LET } Q1b \triangleq \{m \in msgs : \land m.type = ``1b'' \\ \land m.acc \in Q \\ \land m.bal = b \} \\ Q1bv \triangleq \{m \in Q1b : m.mbal \geq 0 \} \\ \text{IN } \land \forall \ a \in Q : \exists \ m \in Q1b : m.acc = a \\ \land \lor Q1bv = \{\} \\ \lor \exists \ m \in Q1bv : \\ \land m.mval = v \\ \land \forall \ mm \in Q1bv : m.mbal \geq mm.mbal \end{array}$$

This condition implies that v is safe at bbecause this condition implies ShowsSafeAt(Q, b, v).

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = "1b" \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a$
 $\land \lor Q1bv = \{\}$
 $\lor \exists m \in Q1bv :$
 $\land m.mval = v$
 $\land \forall mm \in Q1bv : m.mbal > mm.mbal$

This implies ShowsSafeAt(Q, b, v)

This implies ShowsSafeAt(Q, b, v)

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ \land \forall a \in Q : maxBal[a] \geq b \\ \land \exists c \in -1 \dots (b-1) : \\ \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

This implies ShowsSafeAt(Q, b, v) where these are defined in terms of messages that have been sent.

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ \land \forall a \in Q : maxBal[a] \geq b \\ \land \exists c \in -1 \dots (b-1) : \\ \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v)] \\ \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = "1b" \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a$
 $\land \lor Q1bv = \{\}$
 $\lor \exists m \in Q1bv :$
 $\land m.mval = v$
 $\land \forall mm \in Q1bv : m.mbal > mm.mbal$

$$ShowsSafeAt(Q, b, v) \triangleq$$

$$\land \forall a \in Q : maxBal[a] \ge b$$

$$\land \exists c \in -1 .. (b-1) :$$

$$\land (c \ne -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v)$$

$$\land \forall d \in (c+1) .. (b-1), a \in Q : DidNotVoteAt(a, d)$$
For every a in Q, maxBal[a] = b was true right after it sent its message.

$$ShowsSafeAt(Q, b, v) \triangleq \\ \land \forall a \in Q : maxBal[a] \ge b \\ \land \exists c \in -1 .. (b-1) : \\ \land (c \ne -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ \land \forall d \in (c+1) .. (b-1), a \in Q : DidNotVoteAt(a, d) \end{cases}$$

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = \text{``1b''} \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a \land \lor Q1bv = \{\} \lor \exists m \in Q1bv : \land m.mval = v \land \forall mm \in Q1bv : m.mbal \ge mm.mbal$

And maxBal[a] never decreases.

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ \wedge \overleftarrow{\forall a \in Q : maxBal[a] \ge b} \\ \wedge \exists c \in -1 .. (b-1) : \\ \wedge (c \ne -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ \wedge \forall d \in (c+1) .. (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = "1b" \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a$
 $\land \lor Q1bv = \{\}$
 $\lor \exists m \in Q1bv :$
 $\land m.mval = v$
 $\land \forall mm \in Q1bv : m.mbal > mm.mbal$

$$\begin{array}{l} ShowsSafeAt(Q, \ b, \ v) \triangleq \\ \land \forall \ a \in Q : maxBal[a] \ge b \\ \land \exists \ c \in -1 \dots (b-1) : \\ \land \ (c \neq -1) \Rightarrow \exists \ a \in Q : VotedFor(a, \ c, \ v) \\ \land \forall \ d \in (c+1) \dots (b-1), \ a \in Q : DidNotVoteAt(a, \ d) \end{array}$$

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = "1b" \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a$
 $\land \lor Q1bv = \{\}$
 $\lor \exists m \in Q1bv :$
 $\land m.mval = v$
 $\land \forall mm \in Q1bv : m.mbal > mm.mbal$

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ & \land \forall \ a \in Q : maxBal[a] \geq b \\ & \land \exists \ c \in -1 \dots (b-1) : \\ & \land (c \neq -1) \Rightarrow \exists \ a \in Q : VotedFor(a, c, v) \\ & \land \forall \ d \in (c+1) \dots (b-1), \ a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = "1b" \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a$
 $\land \lor Q1bv = \{\}$
 $\lor \exists m \in Q1bv :$
 $\land m.mval = v$
 $\land \forall mm \in Q1bv : m.mbal > mm.mbal$

When an acceptor $a \in Q$ sent its 1 *b* message mm for ballot *b*,

And because it set mBal[a] to b, it still hasn't voted in those ballots.

And because it set mBal[a] to b, it still hasn't voted in those ballots.

$$\begin{array}{ll} & \wedge \forall \, a \in Q : \exists \, m \in Q1b : m.acc = a \\ & \wedge \lor Q1bv = \{\} \\ & \lor \exists \, m \in Q1bv : \\ & \land m.mval = v \\ & \land \forall \, mm \in Q1bv : m.mbal \geq mm.mbal \end{array}$$

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = "1b" \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a$
 $\land \lor Q1bv = \{\}$
 $\lor \exists m \in Q1bv :$
 $\land m.mval = v$
 $\land \forall mm \in Q1bv : m.mbal > mm.mbal$

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = "1b" \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a$
 $\land \lor Q1bv = \{\}$
 $\lor \exists m \in Q1bv :$
 $\land m.mval = v$
 $\land \forall mm \in Q1bv : m.mbal > mm.mbal$

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ & \land \forall a \in Q : maxBal[a] \geq b \\ & \land \exists c \in -1 \dots (b-1) : \\ & \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ & \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = "1b" \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a$
 $\land \lor Q1bv = \{\}$
 $\lor \exists m \in Q1bv :$
 $\land m.mval = v$
 $\land \forall mm \in Q1bv : m.mbal > mm.mbal$

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ & \land \forall \ a \in Q : maxBal[a] \geq b \\ & \land \exists \ c \in -1 \dots (b-1) : \\ & \land (c \neq -1) \Rightarrow \exists \ a \in Q : VotedFor(a, c, v) \\ & \land \forall \ d \in (c+1) \dots (b-1), \ a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = "1b" \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a$
 $\land \lor Q1bv = \{\}$
 $\lor \exists m \in Q1bv :$
 $\land m.mval = v$
 $\land \forall mm \in Q1bv : m.mbal \ge mm.mbal$

Case 1: No acceptor in *Q* has voted.

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ \land \forall \ a \in Q : maxBal[a] \geq b \\ \land \exists \ c \in -1 \dots (b-1) : \\ \land \ (c \neq -1) \Rightarrow \exists \ a \in Q : VotedFor(a, c, v) \\ \land \forall \ d \in (c+1) \dots (b-1), \ a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = "1b" \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a$
 $\land \lor Q1bv = \{\}$
 $\lor \exists m \in Q1bv :$
 $\land m.mval = v$
 $\land \forall mm \in Q1bv : m.mbal \ge mm.mbal$

Let c equal -1.

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ &\land \forall \ a \in Q : maxBal[a] \geq b \\ &\land \exists \ c \in -1 \dots (b-1) : \\ &\land (c \neq -1) \Rightarrow \exists \ a \in Q : VotedFor(a, c, v) \\ &\land \forall \ d \in (c+1) \dots (b-1), \ a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = "1b" \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a$
 $\land \lor Q1bv = \{\}$
 $\lor \exists m \in Q1bv :$
 $\land m.mval = v$
 $\land \forall mm \in Q1bv : m.mbal \ge mm.mbal$

Let c equal -1. This is vacuously true.

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ & \land \forall a \in Q : maxBal[a] \geq b \\ & \land \exists c \in -1 \dots (b-1) : \\ & \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ & \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = "1b" \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a$
 $\land \lor Q1bv = \{\}$
 $\lor \exists m \in Q1bv :$
 $\land m.mval = v$
 $\land \forall mm \in Q1bv : m.mbal > mm.mbal$

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ & \land \forall a \in Q : maxBal[a] \geq b \\ & \land \exists c \in -1 \dots (b-1) : \\ & \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ & \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = "1b" \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a$
 $\land \lor Q1bv = \{\}$
 $\lor \exists m \in Q1bv :$
 $\land m.mval = v$
 $\land \forall mm \in Q1bv : m.mbal \ge mm.mba$

No acceptor a in Q had voted when it sent its message.

$$\begin{array}{l} ShowsSafeAt(Q, b, v) \triangleq \\ \land \forall a \in Q : maxBal[a] \ge b \\ \land \exists c \in -1 \dots (b-1) : \\ \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{array}$$

No acceptor a in Q had voted when it sent its message. And mBal[a] then equaled b, so it couldn't later have voted in any ballot < b.

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ & \land \forall a \in Q : maxBal[a] \geq b \\ & \land \exists c \in -1 \dots (b-1) : \\ & \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ & \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

And because it set mBal[a] to b, it still hasn't voted in those ballots.

$$\begin{array}{ll} \text{IN} & \wedge \forall \, a \in Q : \exists \, m \in Q1b : m.acc = a \\ & \wedge \lor Q1bv = \{\} \\ & \lor \exists \, m \in Q1bv : \\ & \land \, m.mval = v \\ & \land \forall \, mm \in Q1bv : m.mbal \geq mm.mbal \end{array}$$

No acceptor a in Q had voted when it sent its message. And mBal[a] then equaled b, so it couldn't later have voted in any ballot < b.

$$\begin{array}{l} ShowsSafeAt(Q, b, v) \triangleq \\ \land \forall a \in Q : maxBal[a] \ge b \\ \land \exists c \in -1 \dots (b-1) : \\ \land (c \ne -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{array}$$

Case 2:

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ &\land \forall a \in Q : maxBal[a] \geq b \\ &\land \exists c \in -1 \dots (b-1) : \\ &\land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ &\land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

Case 2: Choose such a message m.

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ & \land \forall a \in Q : maxBal[a] \geq b \\ & \land \exists c \in -1 \dots (b-1) : \\ & \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ & \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ &\land \forall a \in Q : maxBal[a] \geq b \\ &\land \exists c \in -1 \dots (b-1) : \\ &\land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ &\land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ \land \forall a \in Q : maxBal[a] \geq b \\ \land \exists c \in -1 \dots (b-1) : \\ \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ & \land \forall a \in Q : maxBal[a] \geq b \\ & \land \exists c \in -1 \dots (b-1) : \\ & \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ & \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

Let c = m.mbal. VotedFor(m.acc, c, v)

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ &\land \forall a \in Q : maxBal[a] \geq b \\ &\land \exists c \in -1 \dots (b-1) : \\ &\land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ &\land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ & \land \forall a \in Q : maxBal[a] \geq b \\ & \land \exists c \in -1 \dots (b-1) : \\ & \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ & \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ \land \forall a \in Q : maxBal[a] \geq b \\ \land \exists c \in -1 \dots (b-1) : \\ \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ \land \forall a \in Q : maxBal[a] \geq b \\ \land \exists c \in -1 \dots (b-1) : \\ \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

Let c = m.mbal. c is the highest numbered ballot in which any $a \in Q$ voted.

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ &\land \forall a \in Q : maxBal[a] \geq b \\ &\land \exists c \in -1 .. (b-1) : \\ &\land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ &\land \forall d \in (c+1) .. (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

And because it set mBal[a] to b, it still hasn't voted in those ballots.

$$\begin{array}{ll} & \wedge \forall \, a \in Q : \exists \, m \in Q1b : m.acc = a \\ & \wedge \lor Q1bv = \{\} \\ & \lor \exists \, m \in Q1bv : \\ & \land m.mval = v \\ & \land \forall \, mm \in Q1bv : m.mbal \geq mm.mbal \end{array}$$

Let c = m.mbal. c is the highest numbered ballot in which any $a \in Q$ voted. No $a \in Q$ voted in any ballot d with c < d < b.

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ & \land \forall a \in Q : maxBal[a] \geq b \\ & \land \exists c \in -1 \dots (b-1) : \\ & \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ & \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

Let c = m.mbal. c is the highest numbered ballot in which any $a \in Q$ voted. No $a \in Q$ voted in any ballot d with c < d < b.

$$\begin{array}{l} ShowsSafeAt(Q, b, v) \triangleq \\ \land \forall a \in Q : maxBal[a] \ge b \\ \land \exists c \in -1 \dots (b-1) : \\ \land (c \ne -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{array}$$

This is an explanation, not a proof.

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ \land \forall a \in Q : maxBal[a] \geq b \\ \land \exists c \in -1 \dots (b-1) : \\ \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

This is an explanation, not a proof.

There are more rigorous proofs

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ & \land \forall a \in Q : maxBal[a] \geq b \\ & \land \exists c \in -1 \dots (b-1) : \\ & \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ & \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

This is an explanation, not a proof.

There are more rigorous proofs of incorrect algorithms.

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ & \land \forall a \in Q : maxBal[a] \geq b \\ & \land \exists c \in -1 \dots (b-1) : \\ & \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ & \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

LET
$$Q1b \triangleq \{m \in msgs : \land m.type = "1b" \land m.acc \in Q \land m.bal = b\}$$

 $Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\}$
IN $\land \forall a \in Q : \exists m \in Q1b : m.acc = a$
 $\land \lor Q1bv = \{\}$
 $\lor \exists m \in Q1bv :$
 $\land m.mval = v$
 $\land \forall mm \in Q1bv : m.mbal > mm.mbal$

$$\begin{aligned} ShowsSafeAt(Q, b, v) &\triangleq \\ & \land \forall a \in Q : maxBal[a] \geq b \\ & \land \exists c \in -1 \dots (b-1) : \\ & \land (c \neq -1) \Rightarrow \exists a \in Q : VotedFor(a, c, v) \\ & \land \forall d \in (c+1) \dots (b-1), a \in Q : DidNotVoteAt(a, d) \end{aligned}$$

$$\begin{array}{l} \begin{array}{l} Phase2a(b, v) \triangleq \\ \land \neg \exists \ m \in msgs: m.type = ``2a'' \land m.bal = b \\ \land \exists \ Q \in Quorum: \\ & \text{LET} \ Q1b \triangleq \{m \in msgs: \land m.type = ``1b'' \\ & \land m.acc \in Q \\ \land m.bal = b\} \end{array}$$

$$\begin{array}{l} \begin{array}{l} Q1bv \triangleq \{m \in Q1b: m.mbal \geq 0\} \\ \text{IN} \quad \land \forall \ a \in Q: \exists \ m \in Q1b: m.acc = a \\ \land \lor Q1bv = \{\} \\ & \lor \exists \ m \in Q1bv: \\ \land m.mval = v \\ \land \forall \ mm \in Q1bv: m.mbal \geq mm.mbal \end{array}$$
$$\begin{array}{l} Phase2a(b, v) \triangleq \\ \wedge \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \\ \wedge \exists \ Q \in Quorum : \\ \text{LET } Q1b \triangleq \{m \in msgs : \land m.type = ``1b'' \\ & \land m.acc \in Q \\ \wedge m.bal = b\} \\ Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\} \\ \text{IN } \land \forall \ a \in Q : \exists \ m \in Q1b : m.acc = a \\ \land \lor Q1bv = \{\} \\ \lor \exists \ m \in Q1bv : \\ \land m.mval = v \\ \land \forall \ mm \in Q1bv : m.mbal \ge mm.mbal \\ \land Send([type \mapsto ``2a'', \ bal \mapsto b, \ val \mapsto v]) \end{array}$$

$$\begin{array}{l} Phase 2a(b, v) \triangleq \\ \land \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \\ \land \exists \ Q \in Quorum : \\ \text{LET } Q1b \triangleq \{m \in msgs : \land m.type = ``1b'' \\ \land m.acc \in Q \\ \land m.bal = b\} \\ Q1bv \triangleq \{m \in Q1b : m.mbal \geq 0\} \\ \text{IN } \land \forall \ a \in Q : \exists \ m \in Q1b : m.acc = a \\ \land \lor Q1bv = \{\} \\ \lor \exists \ m \in Q1bv : \\ \land m.mval = v \\ \land \forall \ mm \in Q1bv : m.mbal \geq mm.mbal \\ \land Send([type \mapsto ``2a'', \ bal \mapsto b, \ val \mapsto v]) \\ \end{array}$$

$$\begin{array}{l} Phase2a(b, v) \triangleq \\ \land \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \\ \land \exists \ Q \in Quorum : \\ \text{LET } Q1b \triangleq \{m \in msgs : \land m.type = ``1b'' \\ \land m.acc \in Q \\ \land m.bal = b\} \\ Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\} \\ \text{IN } \land \forall \ a \in Q : \exists \ m \in Q1b : m.acc = a \\ \land \lor Q1bv = \{\} \\ \lor \exists \ m \in Q1bv : \\ \land m.mval = v \\ \land \forall \ mm \in Q1bv : m.mbal \ge mm.mbal \\ \land \ Send([type \mapsto ``2a'', \ bal \mapsto b, \ val \mapsto v]) \\ \land \text{ UNCHANGED } \langle maxBal, \ maxVBal, \ maxVal \rangle \end{array}$$

$$\begin{array}{l} Phase2a(b, v) \triangleq \\ \land \neg \exists \ m \in msgs : m.type = ``2a'' \land m.bal = b \\ \land \exists \ Q \in Quorum : \\ \text{LET } Q1b \triangleq \{m \in msgs : \land m.type = ``1b'' \\ \land m.acc \in Q \\ \land m.bal = b\} \\ Q1bv \triangleq \{m \in Q1b : m.mbal \ge 0\} \\ \text{IN } \land \forall \ a \in Q : \exists \ m \in Q1b : m.acc = a \\ \land \lor Q1bv = \{\} \\ \lor \exists \ m \in Q1bv : \\ \land m.mval = v \\ \land \forall \ mm \in Q1bv : m.mbal \ge mm.mbal \\ \land \ Send([type \mapsto ``2a'', \ bal \mapsto b, \ val \mapsto v]) \\ \land \text{ UNCHANGED } \langle maxBal, \ maxVBal, \ maxVal \rangle \\ \text{Leaves all variables except } msgs \ \text{unchanged.} \end{array}$$

If acceptor a has not agreed to participate in a ballot numbered > b, then it votes for v in ballot b.

If acceptor a has not agreed to participate in a ballot numbered > b, then it votes for v in ballot b.

Upon receipt of a 2*a* message *m* for ballot *m*.*bal*, If acceptor *a* has not agreed to participate in a ballot numbered > b, then it votes for *v* in ballot *b*.

Upon receipt of a 2a message m for ballot m.bal, if acceptor a has not agreed to participate in a ballot numbered > m.bal, then it votes for m.val in ballot m.bal.

$$Phase2b(a) \triangleq \\ \exists m \in msgs : \\ \land m.type = "2a"$$

$$Phase2b(a) \triangleq \\ \exists m \in msgs : \\ \land m.type = "2a"$$

$$\begin{array}{l} Phase2b(a) \triangleq \\ \exists \ m \in msgs: \\ \land \ m.type = \ ``2a'' \\ \land \ m.bal \geq maxBal[a] \end{array}$$

$$\begin{array}{l} Phase2b(a) \triangleq \\ \exists \ m \in msgs: \\ \land \ m.type = "2a" \\ \land \ m.bal \geq maxBal[a] \end{array}$$

$$\begin{array}{l} Phase2b(a) \triangleq \\ \exists \ m \in msgs: \\ & \land \ m.type = "2a" \\ & \land \ m.bal \ge maxBal[a] \\ & \land \ maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \end{array}$$

$$\begin{array}{l} Phase2b(a) \triangleq \\ \exists \ m \in msgs: \\ \land \ m.type = ``2a'' \\ \land \ m.bal \ge maxBal[a] \\ \land \ maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \\ \end{array}$$

$$\begin{array}{l} Phase2b(a) \triangleq \\ \exists \ m \in msgs: \\ & \land m.type = ``2a'' \\ & \land m.bal \ge maxBal[a] \\ & \land maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \\ & \land maxVBal' = [maxVBal \ \text{EXCEPT} \ ![a] = m.bal] \end{array}$$

$$\begin{array}{l} Phase2b(a) \triangleq \\ \exists \ m \in msgs: \\ & \land m.type = ``2a'' \\ & \land m.bal \geq maxBal[a] \\ & \land maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \\ & \land maxVBal' = [maxVBal \ \text{EXCEPT} \ ![a] = m.bal] \\ & \text{Set } maxVBal[a] \ \text{to } m.bal \ . \end{array}$$

$$\begin{array}{l} Phase2b(a) \triangleq \\ \exists \ m \in msgs: \\ & \land m.type = ``2a'' \\ & \land m.bal \geq maxBal[a] \\ & \land maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \\ & \land maxVBal' = [maxVBal \ \text{EXCEPT} \ ![a] = m.bal] \\ & \land maxVal' = [maxVal \ \text{EXCEPT} \ ![a] = m.val] \end{array}$$

$$\begin{array}{l} Phase2b(a) \triangleq \\ \exists \ m \in msgs: \\ & \land \ m.type = ``2a'' \\ & \land \ m.bal \geq maxBal[a] \\ & \land \ maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \\ & \land \ maxVBal' = [maxVBal \ \text{EXCEPT} \ ![a] = m.bal] \\ & \land \ maxVBal' = [maxVal \ \text{EXCEPT} \ ![a] = m.val] \\ & \land \ maxVal' = [maxVal \ \text{EXCEPT} \ ![a] = m.val] \\ & \textbf{Set} \ \ maxVal[a] \ \textbf{to} \ \ m.val. \end{array}$$

$$\begin{array}{l} Phase2b(a) \triangleq \\ \exists \ m \in msgs: \\ & \land m.type = ``2a'' \\ & \land m.bal \geq maxBal[a] \\ & \land maxBal' = [maxBal \ \text{EXCEPT} \ ![a] = m.bal] \\ & \land maxVBal' = [maxVBal \ \text{EXCEPT} \ ![a] = m.bal] \\ & \land maxVal' = [maxVal \ \text{EXCEPT} \ ![a] = m.val] \end{array}$$

$$\begin{array}{l} Phase2b(a) \triangleq \\ \exists \ m \in msgs : \\ & \land \ m.type = ``2a'' \\ & \land \ m.bal \geq maxBal[a] \\ & \land \ maxBal' = [maxBal \ \text{EXCEPT } ![a] = m.bal] \\ & \land \ maxVBal' = [maxVBal \ \text{EXCEPT } ![a] = m.bal] \\ & \land \ maxVBal' = [maxVal \ \text{EXCEPT } ![a] = m.val] \\ & \land \ maxVal' = [maxVal \ \text{EXCEPT } ![a] = m.val] \\ & \land \ Send([type \mapsto ``2b'', \ acc \mapsto a, \\ & bal \mapsto m.bal, \ val \mapsto m.val]) \end{array}$$

$$\begin{array}{l} Phase2b(a) \triangleq \\ \exists \ m \in msgs : \\ & \land m.type = ``2a'' \\ & \land m.bal \geq maxBal[a] \\ & \land maxBal' = [maxBal \ \text{EXCEPT } ![a] = m.bal] \\ & \land maxVBal' = [maxVBal \ \text{EXCEPT } ![a] = m.bal] \\ & \land maxVal' = [maxVal \ \text{EXCEPT } ![a] = m.val] \\ & \land Send([type \mapsto ``2b'', \ acc \mapsto a, \\ & bal \mapsto m.bal, \ val \mapsto m.val]) \end{array}$$

$Next \triangleq \forall \exists b \in Ballot : \forall Phase1a(b) \\ \forall \exists v \in Value : Phase2a(b, v) \\ \forall \exists a \in Acceptor : Phase1b(a) \lor Phase2b(a)$

$$Next \stackrel{\Delta}{=} \lor \exists b \in Ballot : \lor Phase1a(b) \\ \lor \exists v \in Value : Phase2a(b, v) \\ \lor \exists a \in Acceptor : Phase1b(a) \lor Phase2b(a) \end{cases}$$

 $Spec \stackrel{\Delta}{=} Init \land \Box[Next]_{vars}$

The Paxos Consensus Algorithm Implements The Voting Algorithm

The parameters of the *Voting* spec are: CONSTANTS *Value* Acceptor Quorum The parameters of the *Voting* spec are:

CONSTANTS Value Acceptor Quorum

VARIABLES maxBal votes

They are implemented in the *Paxos* spec by:

The parameters of the *Voting* spec are:

CONSTANTS Value Acceptor Quorum VARIABLES maxBal votes

They are implemented in the Paxos spec by:CONSTANTSValueAcceptorQuorum

They are implemented in the Paxos spec by: CONSTANTS Value Acceptor Quorum VARIABLE maxBal

They are implemented in the *Paxos* spec by: CONSTANTS *Value* Acceptor Quorum VARIABLE maxBal a defined expression votes

They are implemented in the *Paxos* spec by: CONSTANTS *Value* Acceptor Quorum VARIABLE maxBal a defined expression votes

a defined expression votes

$votes \stackrel{\Delta}{=}$
$votes \stackrel{\Delta}{=} \\ [a \in Acceptor \mapsto]$

.]

$$\begin{array}{l} votes \ \stackrel{\Delta}{=} \\ [a \in Acceptor \mapsto \end{array}$$

A function with domain the set of acceptors.

$votes \stackrel{\Delta}{=} [a \in Acceptor \mapsto]$ The set of votes cast by acceptor a.

A function with domain the set of acceptors.

$votes \triangleq \\ [a \in Acceptor \mapsto \\ \{\langle m.bal, m.val \rangle : m \in$

}]

$votes \triangleq$ $[a \in Acceptor \mapsto \{\langle m.bal, m.val \rangle : m \in \text{The set of } 2b \text{ messages sent by } a.$

}]

$\begin{array}{l} \textit{votes} \ \triangleq \\ [a \in \textit{Acceptor} \mapsto \\ \{\langle m.bal, \ m.val \rangle : m \in \{mm \in msgs : \land mm.type = ``2b'' \\ \land mm.acc = a\}\}] \\ \\ \textbf{The set of } 2b \text{ messages} \\ \textbf{sent by } a . \end{array}$

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

Remember that these substitutions are implied.

Theorem Spec \Rightarrow V!Spec

THEOREM Spec \Rightarrow V!Spec

The model checker can check this theorem.

I won't discuss how the Paxos consensus algorithm is used to build fault-tolerant systems.

I won't discuss how the Paxos consensus algorithm is used to build fault-tolerant systems.

See Paxos Made Simple.

What if you don't win a Turing award?

No.

Will it improve your coding?

Will it improve your coding?

Probably not.

Will it improve your coding?

So why bother?

Will it improve your coding?

So why bother?

Coding should be the easiest, least important part of programming.

Will it improve your coding?

So why bother?

Coding should be the easiest, least important part of programming.

If you're having trouble writing a piece of code, you're doing something wrong.

Rosetta



Rosetta



ESA Spacecraft that explored a comet.

Rosetta



ESA Spacecraft that explored a comet.

Several of its instruments were controlled by the Virtuoso real-time operating system.



2 Springe

The next version of Virtuoso.



Formal Development of a Network-Centric RTOS

Software Engineering for Reliable Embedded Systems

Deringer

The next version of Virtuoso.

Its high-level design is described in TLA⁺ .



Formal Development of a Network-Centric RTOS

Software Engineering for Reliable Embedded Systems

🖉 Springer

The next version of Virtuoso.

Its high-level design is described in TLA⁺.

Here's an email from Eric Verhulst, the head of the development team.



The next version of Virtuoso.

Its high-level design is described in TLA⁺.

Here's an email from Eric Verhulst, the head of the development team.

The [TLA⁺] abstraction helped a lot in coming to a much cleaner architecture



The next version of Virtuoso.

Its high-level design is described in TLA⁺.

Here's an email from Eric Verhulst, the head of the development team.

The [TLA⁺] abstraction helped a lot in coming to a much cleaner architecture (we witnessed first-hand the brainwashing done by years of C programming).



The next version of Virtuoso.

Its high-level design is described in TLA⁺.

Here's an email from Eric Verhulst, the head of the development team.

The [TLA⁺] abstraction helped a lot in coming to a much cleaner architecture (we witnessed first-hand the brainwashing done by years of C programming). One of the results was that the code size is about $10 \times$ less than in [Virtuoso].

You don't produce $10 \times$ less code by better coding.

You don't produce $10 \times$ less code by better coding.

You do it with a cleaner architecture

You don't produce $10 \times$ less code by better coding.

better algorithm You do it with a cleaner architecture

You don't produce $10 \times$ less code by better coding.

You do it with a cleaner architecture, which comes from mathematical thinking.
The [TLA⁺] abstraction helped a lot in coming to a much cleaner architecture (we witnessed first-hand the brainwashing done by years of C programming). One of the results was that the code size is about $10 \times$ less than in [Virtuoso].

You don't produce $10 \times$ less code by better coding.

better algorithm You do it with a cleaner architecture, which comes from mathematical thinking.

It doesn't come from thinking in a programming language.

BY CHRIS NEWCOMBE, TIM RATH, FAN ZHANG, BOGDAN MUNTEANU, MARC BROOKER, AND MICHAEL DEARDEUFF

How Amazon Web Services Uses Formal Methods

SINCE 2011, ENCINEERS at Amazon Web Services (AWS) have used formal specification and model checking to help solve difficult design problems in critical systems. Here, we describe our motivation and experience, what has worked well in our problem domain, and what has not. When discussing personal experience we refer to the authors by their initials.

Åt AWS we strive to build services that are simple for customers to use. External simplicity is built on a hidden substrate of complex distributed systems. Such complex internals are required to achieve high availability while running on cost-efficient infrastructure and cope with relentless business growth. As an example of this growth, in 2006, AWS launched S3, its Simple Storage Service. In the following six years, S3 grew to store one trillion objects.³ Less than a year later it had grown to two trillion objects per second.⁴

S3 is just one of many AWS services that store and process data our customers have entrusted to us. To safeguard that data, the core of each service relies on fault-tolerant distributed algorithms for replication, consistency, concurrency control, auto-scaling, load balancing, and other coordination tasks. There are many such algorithms in the literature, but combining them into a cohesive system is a challenge, as the algorithms must usually be modified to interact properly in a real-world system. In addition, we have found it necessary to invent algorithms of our own. We work hard to avoid unnecessary complexity, but the essential complexity of the task remains high.

Complexity increases the probability of human error in design, code, and operations. Errors in the core of the system could cause loss or corruption of data, or violate other interface contracts on which our customers depend. So, before launching a service. we need to reach extremely high confidence that the core of the system is correct. We have found the standard verification techniques in industry are necessary but not sufficient. We routinely use deep design reviews, code reviews, static code analysis, stress testing, and fault-injection testing but still find that subtle bugs can hide in complex concurrent fault-tolerant systems. One reason they do is that human intuition is poor at estimating the true probability of supposedly "extremely rare" combinations of events in systems operating at a scale of millions of requests per second.

» key insights

- Formal methods find bugs in system designs that cannot be found through any other technique we know of.
- Formal methods are surprisingly feasible for mainstream software development and give good return on investment.
- At Amazon, formal methods are routinely applied to the design of complex real-world software, including public cloud services.

BY CHRIS NEWCOMBE, TIM RATH, FAN ZHANG, BOGDAN MUNTEANU, MARC BROOKER, AND MICHAEL DEARDEUFF

How Amazon Web Services Uses Formal Methods

SINCE 2011, ENCINEERS at Amazon Web Services (AWS) have used formal specification and model checking to help solve difficult design problems in critical systems. Here, we describe our motivation and experience, what has worked well in our problem domain, and what has not. When discussing personal experience we refer to the authors by their initials.

Åt AWS we strive to build services that are simple for customers to use. External simplicity is built on a hidden substrate of complex distributed systems. Such complex internals are required to achieve high availability while running on cost-efficient infrastructure and cope with relentless business growth. As an example of this growth, in 2006, AWS launched S3, its Simple Storage Service. In the following six years, S3 grew to store one trillion objects.³ Less than a year later it had grown to two trillion objects per second.⁴ S3 is just one of many AWS services that store and process data our customers have entrusted to us. To safeguard that data, the core of each service relies on fault-tolerant distributed algorithms for replication, consistency, concurrency control, auto-scaling, load balancing, and other

They build Amazon's cloud infrastructure.

to invent aigorithms or our own. we work hard to avoid unnecessary complexity, but the essential complexity of the task remains high.

Complexity increases the probability of human error in design, code, and operations. Errors in the core of the system could cause loss or corruption of data, or violate other interface contracts on which our customers depend. So, before launching a service. we need to reach extremely high confidence that the core of the system is correct. We have found the standard verification techniques in industry are necessary but not sufficient. We routinely use deep design reviews, code reviews, static code analysis, stress testing, and fault-injection testing but still find that subtle bugs can hide in complex concurrent fault-tolerant systems. One reason they do is that human intuition is poor at estimating the true probability of supposedly "extremely rare" combinations of events in systems operating at a scale of millions of requests per second.

» key insight

- Formal methods find bugs in system designs that cannot be found through any other technique we know of.
- Formal methods are surprisingly feasible for mainstream software development and give good return on investment.
- At Amazon, formal methods are routinely applied to the design of complex real-world software, including public cloud services.

BY CHRIS NEWCOMBE, TIM RATH, FAN ZHANG, BOGDAN MUNTEANU, MARC BROOKER, AND MICHAEL DEARDEUFF

How Amazon Web Services Uses Formal Methods

SINCE 2011, ENCINEERS at Amazon Web Services (AWS) have used formal specification and model checking to help solve difficult design problems in critical systems. Here, we describe our motivation and experience, what has worked well in our problem domain, and what has not. When discussing personal experience we refer to the authors by their initials.

Åt AWS we strive to build services that are simple for customers to use. External simplicity is built on a hidden substrate of complex distributed systems. Such complex internals are required to achieve high availability while running on cost-efficient infrastructure and cope with relentless business growth. As an example of this growth, in 2006, AWS launched S3, its Simple Storage Service. In the following six years, S3 grew to store one trillion objects.³ Less than a year later it had grown to two trillion objects per second.⁴

S3 is just one of many AWS services that store and process data our customers have entrusted to us. To safeguard that data, the core of each service relies on fault-tolerant distributed algorithms for replication, consistency, concurrency control, auto-scaling, load balancing, and other coordination tasks. There are many such algorithms in the literature, but combining them into a cohesive system is a challenge, as the algorithms must usually be modified to interact properly in a real-world system. In addition, we have found it necessary to invent algorithms of our own. We work hard to avoid unnecessary com-

The formal method

they use is TLA⁺.

pend. So, before launching a service. we need to reach extremely high confidence that the core of the system is correct. We have found the standard verification techniques in industry are necessary but not sufficient. We routinely use deep design reviews, code reviews, static code analysis, stress testing, and fault-injection testing but still find that subtle bugs can hide in complex concurrent fault-tolerant systems. One reason they do is that human intuition is poor at estimating the true probability of supposedly "extremely rare" combinations of events in systems operating at a scale of millions of requests per second.

» key insights

- Formal methods find bugs in system designs that cannot be found through any other technique we know of.
- Formal methods are surprisingly feasible for mainstream software development and give good return on investment.
- At Amazon, formal methods are routinely applied to the design of complex real-world software, including public cloud services.

BY CHRIS NEWCOMBE, TIM RATH, FAN ZHANG, BOGDAN MUNTEANU, MARC BROOKER, AND MICHAEL DEARDEUFF

How Amazon Web Services Uses Formal Methods

SINCE 2011, ENCINEERS all Amazon Web Services (AWS) have used formal specification and model checking to help solve difficult design problems in critical systems. Here, we describe our motivation and experience, what has worked well in our problem domain, and what has not. When discussing personal experience we refer to the authors by their initials.

Åt AWS we strive to build services that are simple for customers to use. External simplicity is built on a hidden substrate of complex distributed systems. Such complex internals are required to achieve high availability while running on cost-efficient infrastructure and cope with relentless business growth. As an example of this growth, in 2006, AWS launched S3, its Simple Storage Service. In the following six years, 53 grew to store one trillion objects.³ Less than a year later it had grown to two trillion objects and was regularly handling 1.1 million requests per second.⁴

S3 is just one of many AWS services that store and process data our customers have entrusted to us. To safeguard that data, the core of each service relies on fault-tolerant distributed algorithms for replication, consistency, concurrency control, auto-scaling, load balancing, and other coordination tasks. There are many such algorithms in the literature, but combining them into a cohesive system is a challenge, as the algorithms must usually be modified to interact properly in a real-world system. In addition, we have found it necessary to invent algorithms of our own. We work hard to avoid unnecessary complexity, but the essential complexity of the task remains high.

Complexity increases the probability of human error in design, code, and operations. Errors in the core of the system could cause loss or corruption of data, or violate other interface contracts on which our customers depend. So, before launching a service. we need to reach extremely high confidence that the core of the system is correct. We have found the standard verification techniques in industry are necessary but not sufficient. We routinely use deep design reviews, code reviews, static code analysis, stress testing, and fault-injection testing but still find that subtle bugs can hide in complex concurrent fault-tolerant systems. One reason they do is that human intuition is poor at estimating the true probability of supposedly "extremely rare" combinations of events in systems operating at a scale of millions of requests per second.

» key insights

- Formal methods find bugs in system designs that cannot be found through any other technique we know of.
- Formal methods are surprisingly feasible for mainstream software development and give good return on investment.
- At Amazon, formal methods are routinely applied to the design of complex real-world software, including public cloud services.

BY CHRIS NEWCOMBE, TIM RATH, FAN ZHANG, BOGDAN MUNTEANU, MARC BROOKER, AND MICHAEL DEARDEUFF

How Amazon Web Services Uses Formal Methods

SINCE 2011, ENCINEERS at Amazon Web Services (AWS) have used formal specification and model checking to help solve difficult design problems in critical systems. Here, we describe our motivation and experience, what has worked well in our problem domain, and what has not. When discussing personal experience we refer to the authors by their initials.

Åt AWS we strive to build services that are simple for customers to use. External simplicity is built on a hidden substrate of complex distributed systems. Such complex internals are required to achieve high availability while running on cost-efficient infrastructure and cope with relentless business growth. As an example of this growth, in 2006, AWS launched S3, its Simple Storage Service. In the following six years, 53 grew to store one trillion objects.³ Less than a year later it had grown to two trillion objects and was regularly handling 1.1 million requests per second.⁴

customers ha safeguard that service relies tributed algo consistency, o to-scaling, loa coordination such algorithm combining the tem is a challe must usually properly in a addition, we to invent algo work hard to plexity, but th the task remain Complexity ity of human and operation the system con tion of data, o contracts on pend, So, bef we need to re fidence that t correct. We h verification te necessary but tinely use de reviews, stati testing, and fa

S3 is just

vices that stor

still find that subtle bugs can hide in complex concurrent fault-tolerant systems. One reason they do is that human intuition is poor at estimating the true probability of supposedly "extremely rare" combinations of events in systems operating at a scale of millions of recursets per second.

» key insights

- Formal methods find bugs in system designs that cannot be found through any other technique we know of.
- Formal methods are surprisingly feasible for mainstream software development and give good return on investment.
- At Amazon, formal methods are routinely applied to the design of complex real-world software, including public cloud services.

key insights

Formal methods find bugs in system designs that cannot be found through any other technique we know of.

BY CHRIS NEWCOMBE, TIM RATH, FAN ZHANG, BOGDAN MUNTEANU, MARC BROOKER, AND MICHAEL DEARDEUFF

How Amazon Web Services Uses Formal Methods

SINCE 2011, ENCINEERS all Amazon Web Services (AWS) have used formal specification and model checking to help solve difficult design problems in critical systems. Here, we describe our motivation and experience, what has worked well in our problem domain, and what has not. When discussing personal experience we refer to the authors by their initials.

Åt AWS we strive to build services that are simple for customers to use. External simplicity is built on a hidden substrate of complex distributed systems. Such complex internals are required to achieve high availability while running on cost-efficient infrastructure and cope with relentless business growth. As an example of this growth, in 2006, AWS launched S3, its Simple Storage Service. In the following six years, 53 grew to store one trillion objects.³ Less than a year later it had grown to two trillion objects and was regularly handling 1.1 million requests per second.⁴

safeguard that service relies tributed algoconsistency, o to-scaling, los coordination such algorithm combining th tem is a chall must usually properly in addition, we to invent algo work hard to plexity, but th the task remain Complexity ity of human and operation the system con tion of data, o contracts on pend, So, bef we need to re fidence that t correct. We h verification te necessary but tinely use de reviews, stati testing, and fa

S3 is just

vices that stor

customers ha

still find that subtle bugs can hide in complex concurrent fault-tolerant systems. One reason they do is that human intuition is poor at estimating the true probability of supposedly "extremely rare" combinations of events in systems operating at a scale of millions of recursets per second.

» key insights

- Formal methods find bugs in system designs that cannot be found through any other technique we know of.
- Formal methods are surprisingly feasible for mainstream software development and give good return on investment.
- At Amazon, formal methods are routinely applied to the design of complex real-world software, including public cloud services.

» key insights

Formal methods are surprisingly feasible for mainstream software development and give good return on investment.

BY CHRIS NEWCOMBE, TIM RATH, FAN ZHANG, BOGDAN MUNTEANU, MARC BROOKER, AND MICHAEL DEARDEUFF

How Amazon Web Services Uses Formal Methods

SINCE 2011, ENCINEERS at Amazon Web Services (AWS) have used formal specification and model checking to help solve difficult design problems in critical systems. Here, we describe our motivation and experience, what has worked well in our problem domain, and what has not. When discussing personal experience we refer to the authors by their initials.

Åt AWS we strive to build services that are simple for customers to use. External simplicity is built on a hidden substrate of complex distributed systems. Such complex internals are required to achieve high availability while running on cost-efficient infrastructure and cope with relentless business growth. As an example of this growth, in 2006, AWS launched S3, its Simple Storage Service. In the following six years, 53 grew to store one trillion objects.³ Less than a year later it had grown to two trillion objects and was regularly handling 1.1 million requests per second.⁴

vices that stor customers ha safeguard that service relies tributed algoconsistency, o to-scaling, los coordination such algorithm combining th tem is a chall must usually properly in addition, we to invent algo work hard to plexity, but th the task remain Complexity ity of human and operation the system con tion of data, o contracts on v pend. So, befo we need to re fidence that t correct. We h verification tec necessary but tinely use dee reviews, stati testing, and fa

S3 is just

still find that subtle bugs can hide in complex concurrent fault-tolerant systems. One reason they do is that human intuition is poor at estimating the true probability of supposedly "extremely rare" combinations of events in systems operating at a scale of millions of recursets per second.

» key insights

- Formal methods find bugs in system designs that cannot be found through any other technique we know of.
- Formal methods are surprisingly feasible for mainstream software development and give good return on investment.
- At Amazon, formal methods are routinely applied to the design of complex real-world software, including public cloud services.

key insights

At Amazon, formal methods are routinely applied to the design of complex real-world software, including public cloud services.

66 COMMUNICATIONS OF THE ACH | APRIL 2015 | VOL. 58 | NO. 4

BY CHRIS NEWCOMBE, TIM RATH, FAN ZHANG, BOGDAN MUNTEANU, MARC BROOKER, AND MICHAEL DEARDEUFF

How Amazon Web Services Uses Formal Methods

SINCE 2011, ENCINEERS at Amazon Web Services (AWS) have used formal specification and model checking to help solve difficult design problems in critical systems. Here, we describe our motivation and experience, what has worked well in our problem domain, and what has not. When discussing personal experience we refer to the authors by their initials.

At AWS we strive to build services that are simple for customers to use. External simplicity is built on a hidden substrate of complex distributed systems. Such complex internals are required to achieve high availability while running on cost-efficient infrastructure and cope with relentless business growth. As an example of this growth, in 2006, AWS launched S3, its simple Storage Service. In the following six years, S3 grew to store one

S3 is just one of many AWS services that store and process data our customers have entrusted to us. To safeguard that data, the core of each service relies on fault-tolerant distributed algorithms for replication, consistency, concurrency control, auto-scaling, load balancing, and other coordination tasks. There are many such algorithms in the literature, but combining them into a cohesive system is a challenge, as the algorithms must usually be modified to interact properly in a real-world system. In addition, we have found it necessary to invent algorithms of our own. We work hard to avoid unnecessary complexity, but the essential complexity of the task remains high.

Complexity increases the probability of human error in design, code, and operations. Errors in the core of the system could cause loss or corruption of data, or violate other interface contracts on which our customers depend. So, before launching a service. we need to reach extremely high confidence that the core of the system is correct. We have found the standard verification techniques in industry are necessary but not sufficient. We routinely use deep design reviews, code reviews, static code analysis, stress testing, and fault-injection testing but still find that subtle bugs can hide in complex concurrent fault-tolerant systems. One reason they do is that human intuition is poor at estimating the true probability of supposedly "extremely rare" combinations of events in systems operating at a scale of millions of requests per second.

» key insight

- Formal methods find bugs in system designs that cannot be found through any other technique we know of.
- Formal methods are surprisingly feasible for mainstream software development and also good return on Investment.

66 COMMUNICATIONS OF THE ACM | APRIL 2015 | VOL. 58 | NO. 4

TLA+ is the most valuable thing that I've learned in my professional career.

TLA+ is the most valuable thing that I've learned in my professional career.

It has changed how I work

TLA+ is the most valuable thing that I've learned in my professional career.

It has changed how I work by giving me an immensely powerful tool to find subtle flaws in system designs.

TLA+ is the most valuable thing that I've learned in my professional career.

It has changed how I work

It has changed how I think

TLA+ is the most valuable thing that I've learned in my professional career.

It has changed how I work

It has changed how I think by giving me a framework for constructing new kinds of mental-models

TLA+ is the most valuable thing that I've learned in my professional career.

It has changed how I work

It has changed how I think by giving me a framework for constructing new kinds of mental-models by revealing the precise relationship between correctness properties and system designs

TLA+ is the most valuable thing that I've learned in my professional career.

It has changed how I work

It has changed how I think

by giving me a framework for constructing new kinds of mental-models by revealing the precise relationship between correctness properties and system designs and by allowing me to move from 'plausible prose' to precise statements much earlier in the software development process.

TLA+ is the most valuable thing that I've learned in my professional career.

It has changed how I work

It has changed how I think

Mathematical thinking TLA+ is the most valuable thing that I've learned in my professional career.

It has changed how I work

It has changed how I think