Verifying Programs That Use
Causally-Ordered Message-Passing

Scott D. Stoller*
Fred B. Schneider**

TR 94-1423
May 1994

Department of Computer Science
Cornell University
Ithaca, NY 14853-7501

* This author is supported by an IBM Graduate Fellowship.
** This author is supported in part by the Office of Naval Research under contract
N00014-91-J-1219, the National Science Foundation under Grant CCR-9003440,
DARPA/NSF Grant No. CCR-9014363, NASA/DARPA grant NAG-2-893, and AFOSR
grant F49620-94-1-0198. Any opinions, findings, and conclusions or
recommendations expressed in this publication are those of the author and do not
reflect the views of these agencies.
Verifying Programs that use Causally-ordered Message-passing

Scott D. Stoller*         Fred B. Schneider†
Department of Computer Science
Cornell University
Ithaca, New York 14853
May 15, 1994

Abstract

We give an operational model of causally-ordered message-passing primitives. Based on this model, we formulate a Hoare-style proof system for causally-ordered delivery. To illustrate the use of this proof system and to demonstrate the feasibility of applying invariant-based verification techniques to algorithms that depend on causally-ordered delivery, we verify an asynchronous variant of the distributed termination detection algorithm of Dijkstra, Feijen, and van Gasteren.

*This author is supported by an IBM Graduate Fellowship.
†This author is supported in part by the Office of Naval Research under contract N00014-91-J-1219, the National Science Foundation under Grant CCR-9003440, DARPA/NSF Grant No. CCR-9014363, NASA/DARPA grant NAG-2-893, and AFOSR grant F49620-94-1-0198. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not reflect the views of these agencies.
1 Introduction

Causally-ordered delivery can be understood as a generalization of FIFO ordering [vR93]. In both, a message is delivered only after all messages on which it may depend. With FIFO ordering, this guarantee applies only to messages having the same sender; with causal ordering, this guarantee applies to messages sent by any process. Additional motivation for and examples of the use of causally-ordered delivery can be found in [Bir93, vR93].

This paper gives a proof system for causally-ordered delivery. Our proof system is similar in style to the satisfaction-based logics for synchronous message-passing in [LG81], for ordinary asynchronous message-passing in [SS84], and for flush channels in [CKA93]. We assume familiarity with the terminology of that literature.

Reasoning about message-passing primitives for causally-ordered delivery involves a global property: the system-wide causality relation, which defines what messages are deliverable. This distinguishes causally-ordered delivery from the types of message passing for which axiomatic semantics have already been given (e.g., [LG81, SS84, CKA93]). And, our work demonstrates that substantially new methods are not required when message-delivery semantics depends on global information.

A program proof in a satisfaction-based logic involves discharging three obligations:

1. a proof outline characterizes execution of each process in isolation,

2. a "satisfaction proof" validates postconditions of receive statements, and

3. an interference-freedom proof establishes that execution of no process invalidates an assertion in another.

Our proof system for causally-ordered message-passing is similar, except step (2) is merged with step (1). (Such a merging is also possible for other satisfaction-based proof systems that handle asynchronous communication primitives, like the logics of [SS84] and [CKA93].)

The remainder of the paper is organized as follows. Section 2 defines causally-ordered message-passing. Our proof system is the subject of Section 3. In Section 4, we use the proof system to verify an asynchronous variant of the distributed termination detection algorithm of Dijkstra, Feijen, and van Gasteren [DFvG83]. Section 5 contains some conclusions.
2 A Model of Causally-ordered Message-passing

We give an operational semantics for causally-ordered message-passing primitives by translating programs containing these primitives into a generic concurrent programming language that has shared variables. The shared variables represent the state of the network.

Processes communicate by sending and receiving messages. To encode the restrictions implicit in causally-ordered delivery, each message sent is modeled in our translation by a triple \( (d, i, t) \), where\(^1\)

\[
\begin{align*}
d & \text{ is the data being sent by the program,} \\
i & \text{ is the name of the process\(^2\) that sent the message, and} \\
t & \text{ is a timestamp that contains information used to determine whether the message is ready for delivery.}
\end{align*}
\]

The following functions are useful in connection with messages represented by triples.

\[
\begin{align*}
data((d, i, t)) & \triangleq d \\
sender((d, i, t)) & \triangleq i \\
ts((d, i, t)) & \triangleq t
\end{align*}
\]

Two shared variables \( \sigma_i \) and \( \rho_i \) are associated with each process \( i \). Variable \( \sigma_i \) contains the (triples modeling) messages sent to process \( i \); \( \rho_i \) contains the (triples modeling) messages process \( i \) has received.

There is an obvious and seemingly simpler alternative to using variables \( \sigma_i \) and \( \rho_i \). It is to use a single variable \( \chi_i \) (say), where the value of \( \chi_i \) is the set of messages sent to process \( i \) but not yet received (i.e., \( \chi_i \) equals \( \sigma_i - \rho_i \)). The model we use has two advantages over this one-variable model. First, in our model, proving interference freedom (defined in Section 3) is easier. This is because no process can falsify \( m \in \sigma_i \) or \( m \in \rho_i \); predicate \( m \in \chi_i \) would be invalidated by the receiver. Second, proofs of some programs (such as the example in Section 4) involve reasoning about communications history. That history is available in \( \sigma_i \) and \( \rho_i \) but is not available in \( \chi_i \).

Causally-ordered delivery restricts when a message can be received. This is achieved in our

\(^1\) An actual implementation of causally-ordered delivery might not require a sender name \( i \) or timestamp \( t \). That information is used here to abstract from the details of all real implementations.

\(^2\) Processes are named \( 0, 1, \ldots, N - 1 \), and hereafter identifiers \( i, j, k, \) and \( p \) range over process names.
translation by defining a well-founded partial order \(<\) on timestamps. Our definition of \(<\) is based on the theory of [Lam78]. A system execution is represented as a tuple of sequences of events; each sequence corresponds to the execution of a single process. An event is a send event, a receive event, or an internal (i.e., non-communication) event. The happens-before (or “potential causality”) relation \(\rightarrow\) for a system execution is the smallest transitive binary relation on the events in that execution such that:

- If \(e\) and \(e'\) are performed by the same process and \(e\) occurs before \(e'\), then \(e \rightarrow e'\).
- If \(e\) is the send event for a message \(m\) and \(e'\) is the receive event for that message, then \(e \rightarrow e'\).

Causally-ordered delivery is formalized in terms of \(\rightarrow\) as follows [BSS91]. Let \(\text{send}(m)\) and \(\text{receive}(m)\) respectively denote the send event and receive event for a message \(m\).

**Causally-ordered Delivery:** If \(m\) and \(m'\) are sent to the same process and \(\text{send}(m) \rightarrow \text{send}(m')\), then \(\text{receive}(m) \rightarrow \text{receive}(m')\).\(^3\)

To implement Causally-ordered Delivery using timestamped messages, the timestamps and \(<\) are chosen to satisfy

\[
ts(m) < ts(m') \iff \text{send}(m) \rightarrow \text{send}(m').
\]

Causally-ordered Delivery is then equivalent to requiring that a message \(m'\) is received by a process \(p\) only after \(p\) has received all messages \(m\) sent to \(p\) for which \(ts(m) < ts(m')\) holds.

One way to achieve (1) is to use vector clocks [Fid88, Mat89]. Here, a vector \(vt_i\) of type array[0..\(N - 1\)] of Nat is associated with process \(i\), where \(vt_i\) satisfies:

**Vector Clock Property:** \(vt_i[j]\) is the number of send events that are performed by process \(j\) and causally precede the next event to be performed by process \(i\).

Partial order \(<\) is defined in terms of vector clocks, as follows.

\[
vt_1 \neq vt_2 \triangleq (\exists i : vt_1[i] \neq vt_2[i])
\]
\[
vt_1 < vt_2 \triangleq (\forall i : vt_1[i] \leq vt_2[i]) \land vt_1 \neq vt_2
\]

\(^3\)FIFO delivery can also be formalized in terms of \(\rightarrow\). FIFO delivery ensures that if \(m\) and \(m'\) are sent by the same process, to the same process, and \(\text{send}(m) \rightarrow \text{send}(m')\), then \(\text{receive}(m) \rightarrow \text{receive}(m')\). The close analogy between FIFO delivery and causally-ordered delivery should now be evident.
Three rules define how the $vt_i$ are updated in order to maintain the Vector Clock Property. Since only send events and receive events are of interest, vector clocks are updated only when send and receive statements are executed. Let $inc(vt, i)$ denote vector $vt$ with the $i^{th}$ component incremented by one. The rules are:

**Initialization Rule:** Initially, $vt_i[j] = 0$ for all $i$ and $j$.

**Send Update Rule:** When process $i$ sends a message $m$, it updates $vt_i$ by executing

$$vt_i := inc(vt_i, i)$$

and includes updated vector $vt_i$ as the timestamp attached to $m$.

**Receive Update Rule:** When a process $i$ receives a message $m$, it updates $vt_i$ by executing

$$vt_i := \max(vt_i, ts(m)),$$

where $\max(vt, vt')$ is the component-wise maximum of the vectors $vt$ and $vt'$.

We now give our translation of send and receive statements into statements that read and write shared variables $\sigma_i$ and $\rho_i$. The following notation is used to describe the multiple-assignment [Gri76] of $e_1$ to $x_1$, $e_2$ to $x_2$, ..., and $e_n$ to $x_n$:

$$
\begin{pmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
\end{pmatrix} :=
\begin{pmatrix}
  e_1 \\
  e_2 \\
  \vdots \\
  e_n
\end{pmatrix}
$$

A send statement **send** $e$ to $i$ in process $j$ is translated into:

$$
\begin{pmatrix}
  vt_j \\
  \sigma_i
\end{pmatrix} :=
\begin{pmatrix}
  inc(vt_j, j) \\
  \sigma_i \oplus (e, j, inc(vt_j, j))
\end{pmatrix}
$$

(2)

where $s \oplus x \triangleq s \cup \{x\}$.

The translation of a receive statement requires a conditional delay. Statement **await** $B$ then $S$ delays until $B$ holds and then executes $S$ as a single indivisible operation starting from a state that satisfies $B$. A receive statement **receive** $x$ in process $i$ delays until a message is available
for receipt and then updates $x$, $\rho_i$, and $vt_i$. In particular, to ensure causally-ordered delivery, receive $x$ delays until there exists some message $m$ that has been sent to $i$ but not received and such that all messages $m'$ that have been or will be sent to $i$ for which $ts(m') \prec ts(m)$ have been received.

For a set $A$ of triples modeling messages, $\text{choose}(A)$ and $\text{minset}(A)$ are assumed to satisfy

$$\text{choose}(A) \in A \quad \text{provided } A \neq \emptyset$$

$$\text{minset}(A) \triangleq \{ m \in A \mid (\forall m' \in A : \neg (ts(m') \prec ts(m))) \}$$

A receive statement receive $x$ in process $i$ is translated as follows, where $m_i$ is a fresh variable.

$$\text{await } \sigma_i - \rho_i \neq \emptyset \text{ then } m_i := \text{choose}(\text{minset}(\sigma_i - \rho_i))$$

$x := \text{data}(m_i)$

$vt_i := \max(vt_i, ts(m_i))$

$$\rho_i := \rho_i \oplus m_i$$

To show that code fragments (2) and (5) correctly implement Causally-ordered Delivery, consider some message $m$ that is received by a process $i$. We must show that no message $m'$ subsequently received by process $i$ satisfies $\text{send}(m') \rightarrow \text{send}(m)$. Suppose such a message $m'$ exists. By (1), $ts(m') \prec ts(m)$. Message $m'$ could not be in $\sigma_i$ when $m$ is received, since $m$ is selected from among the elements of $\sigma_i$ with minimal timestamps. Thus, $m'$ must be added to $\sigma_i$ after $m$ is received. We show that this is impossible by proving: For all messages $m$ and $m'$, if $m'$ is added to $\sigma_i$ after $m$ has been added, then $\neg(ts(m') \prec ts(m))$.

First, observe that the following holds throughout execution of a program.

$$(\forall j, k : \: vt_j[k] \leq vt_k[k] \land (\forall m \in \sigma_j : \: ts(m)[k] \leq vt_k[k]))$$

Initially, this holds because for all $j$ and $k$, $vt_j[k] = 0$ and $\sigma_j = \emptyset$. Only send and receive statements change the values of these variables, so it suffices to show that our translations of these statements preserve (6), which is easily done.

Finally, we show that $\neg(ts(m') \prec ts(m))$. This is implied by $\exists k : \: ts(m')[k] > ts(m)[k]$, which, in turn, follows from $ts(m')[j] > ts(m)[j]$ where $j$ is the sender of $m'$. The latter holds because $ts(m')[j] = vt_j[j] + 1 > vt_j[j] \geq ts(m)[j]$, where the equality follows from the translation of send
statements, the strict inequality follows from standard arithmetic, and the nonstrict inequality follows from (6).

3 Axioms for Send and Receive

We can now present Hoare-style axioms [Hoa78] for the send and receive statements described above.

Given the above translation of send e to i into a multiple-assignment statement, we use the multiple-assignment axiom [Gri76] to obtain an axiom for the send statement. The notation e[x_1 := e_1, \ldots, x_n := e_n] denotes the simultaneous substitution of each term e_i for the corresponding variable x_i in a term e. Validity of the following triple follows immediately from the multiple-assignment axiom:

\[
\begin{align*}
\{ & P[vt_j := inc(vt_j, j), \sigma_i := \sigma_i \oplus \langle e, j, inc(vt_j, j) \rangle] \\
( & vt_j) := \left( \begin{array}{c} inc(vt_j, j) \\
\sigma_i \oplus \langle e, j, inc(vt_j, j) \rangle \end{array} \right) \\
\{ & P \}
\end{align*}
\]

Thus, we have

**Send Axiom:** For a send statement in process j:

\[
\{ P[vt_j := inc(vt_j, j), \sigma_i := \sigma_i \oplus \langle e, j, inc(vt_j, j) \rangle] \} \textbf{ send } e \textbf{ to } i \{ P \} \quad (7)
\]

An inference rule for receive statements is obtained using translation (5) of receive x. Using axiom (3) for choose, the usual rules for assignment and sequential composition, and this inference rule for await statements [OG76]

**Await Rule:**

\[
\frac{\{ P \land B \} S \{ Q \}}{\{ P \} \textit{ await } B \textit{ then } S \{ Q \}} \quad (8)
\]

we can show that \{ P \} receive x \{ Q \} is valid iff the following Predicate Logic formula is valid:

\[
P \land m_i \in \text{minset}(\sigma_i - \rho_i) \\
\Rightarrow Q[x := \text{data}(m_i), vt_i := \text{max}(vt_i, ts(m_i)), \rho_i := \rho_i \oplus m_i].
\]

Thus, the inference rule for receive statements is
**Receive Rule:** For a receive statement in process \( j \):

\[
P \land m_i \in \text{minset}(\sigma_i - \rho_i) \\
\Rightarrow Q[x := \text{data}(m_i), vt_i := \max(vt_i, ts(m_i)), \rho_i := \rho_i \oplus m_i] \\
\{P\} \text{receive } x \{Q\}
\]

(9)

**Interference Freedom**

The preceding rules for send and receive, together with rules for other statements and the usual miscellaneous rules of Hoare logics (e.g., the Rule of Consequence), can be used to construct a proof outline for each process in isolation. A proof outline is a program annotated with an assertion before and after every statement. A proof outline characterizes the behavior of a process assuming that no other process invalidates assertions in that proof outline. The proof outlines for processes that execute concurrently are combined to obtain a proof outline for the entire system by showing interference freedom [OG76]— that no process invalidates assertions in the proof outline of another process.

In a proof outline \( PO \), the assertion that precedes a statement \( S \) is called the *precondition* of \( S \) and is denoted \( \text{pre}(S) \), the assertion that follows a statement \( S \) is called the *postcondition* of \( S \) and is denoted \( \text{post}(S) \), and we write \( \text{pre}(PO) \) and \( \text{post}(PO) \) to denote the first and last assertions, respectively, in \( PO \). We write \( \{P\} \text{PO} \{Q\} \) to denote the triple obtained by changing \( \text{pre}(PO) \) to \( P \) and \( \text{post}(PO) \) to \( Q \).

An assertion \( P \) appearing in a proof outline \( PO_i \) is *interference free* with respect to proof outlines \( PO_1, \ldots, PO_N \) if for all assignments, sends, and receives \( S \) in a different proof outline than \( P \),

\[
\{P \land \text{pre}(S)\} S \{P\}
\]

(10)

is valid. This is because (10) asserts that execution of \( S \) does not invalidate \( P \). Assignment to variables is the only way to invalidate an assertion.\(^4\) Since our translations for send and receive contain assignments, the interference freedom obligations require checking (10) for each send and receive statement, as well as for each assignment to an ordinary program variable.

Proof outlines \( PO_1, \ldots, PO_N \) are interference free if all assertions \( P \) in the proof outlines are interference free in \( PO_1, \ldots, PO_N \). This leads to the following inference rule.

\(^4\)This is actually an assumption about the assertion language. For example, it rules out allowing control predicates in assertions.
Parallel Composition Rule:

\[
\begin{align*}
\frac{PO_1, \ldots, PO_N \quad PO_1, \ldots, PO_N \text{ are interference free}}{\{\Pi_i \text{ pre}(PO_i)\}[PO_1 \parallel \cdots \parallel PO_N]\{\Pi_i \text{ post}(PO_i)\}}
\end{align*}
\]  

(11)

Note that, in contrast to the logics for asynchronous communication in [SS84] and [CKA93], our parallel composition rule does not have a "satisfaction" obligation. This is not an artifact of causally-ordered message-passing; the logics of [SS84] and [CKA93] could be similarly formulated.

4 Example: Distributed Termination Detection

To illustrate our proof rules, we give a proof outline for the termination detection algorithm of [DFvG83]. Validity of this proof outline shows that the algorithm correctly detects quiescence in systems of processes that communicate using causally-ordered message-passing. Our proof outline is based on the correctness argument given in [DFvG83], modified for causally-ordered delivery instead of the synchronous communication assumed there.\(^5\)

The algorithm is intended for use in systems where processes behave as follows: At each instant, a process is either active or quiescent, where the only action possible by a quiescent process is receipt of a message. A quiescent process may become active upon receipt of a message; an active process becomes quiescent spontaneously. Each process \(i\) has the form

\[
\begin{align*}
&\text{\textit{Init}}_i \\
do& \\
&\begin{cases}
  \quad g_{ij} \rightarrow \text{\textit{send}} \; e_{ij} \; \text{to} \; j \\
  \quad S_{ij} \\
  \quad \text{\textit{receive}} \; x_i \rightarrow R_i
\end{cases} \\
&\text{\textit{od}}
\end{align*}
\]  

(12)

where the \(g_{ij}\) are boolean expressions, and \text{\textit{Init}}_i, S_{ij}, and \(R_i\) are statements that do not contain communication statements. Such a process \(i\) is quiescent iff each guard \(g_{ij}\) is false. This is formalized by:

\[
q_i \triangleq \neg (\bigvee_j g_{ij})
\]

In the algorithm of [DFvG83] a token circulates among the processes. This introduces a new kind of message, which we call a token message. To distinguish it from the messages in the original

\(^5\)In [Apt86], the partial-correctness argument of [DFvG83] is formalized and some additional properties of the algorithm are proven.
computation, hereafter called basic messages, we use a predicate $istok(data(m))$ that holds exactly
when $m$ is a token message. Note that a process of the form (12) cannot send basic messages to
itself. Define:

\[
\begin{align*}
\sigma_i^{tok} & \triangleq \{ m \in \sigma_i \mid istok(data(m)) \} \\
\rho_i^{tok} & \triangleq \{ m \in \rho_i \mid istok(data(m)) \} \\
\chi_{i,j} & \triangleq \{ m \in \sigma_j - \rho_j \mid \neg istok(data(m)) \land sender(m) = i \}
\end{align*}
\]

The system is quiescent if every process is quiescent and no messages are in transit. Thus, the
system is quiescent iff the following predicate $Q$ holds.

\[
Q \triangleq (\forall i : q_i \land (\forall j : \chi_{i,j} = \emptyset))
\]

A color, either black or white, is associated with each process. For each process $i$, we introduce
a boolean variable $b_i$ such that $b_i$ is true iff process $i$ is black. The detection algorithm sets $b_i$
to true when process $i$ sends a basic message; its sets $b_i$ to false when $i$ sends a token message.
Therefore, we can assert that $b_i$ holds if process $i$ has a sent a basic message since it last sent a
token message. This is formalized as an assertion in terms of the following state function:$^7$

\[
l_{x_i} : \text{The largest timestamp in } \{ m \in \bigcup_j \sigma_j^{tok} \mid sender(m) = i \}, \text{if such a timestamp exists;}
\]
\[
\text{otherwise } 0.
\]

The assertion is now formalized as:

\[
J_i \triangleq (\forall i : (\exists j : (\exists m \in \chi_{i,j} : l_{x_i} < ts(m) )) \Rightarrow b_i)
\]

The algorithm proceeds as a sequence of rounds. One process serves as the initiator for all
rounds; it starts each round by sending a token message. Without loss of generality, assume
process 0 is the initiator. In each round, the token is received by every process exactly once, ending
with the initiator. We define the token to be at position $i$ if it has been sent to process $i$ and not
subsequently sent by process $i$; we say that the token visits a process when the token has been
received by but not sent from that process. For each process $i$, we introduce a new variable $h_i$ that

$^6$This restriction is not needed for correctness of the algorithm; we adopt it here because simplifies the correctness
proof slightly.

$^7$The name $l_{x_i}$ is a mnemonic for “last transmission” of the token by process $i$. 
is true iff the token is visiting process \( i \).

In each round, the token visits the processes in descending order by process name. Thus, the token visits process \( N - 1 \), \( N - 2 \), \ldots, 0, and the current token position is given by the state function:

\[
    tp = \begin{cases} 
        i - 1 & \text{if } \forall j \neq i : l_{x_j} < l_{x_i} \\ 
        N - 1 & \text{otherwise} 
    \end{cases}
\]

Note that all arithmetic on process names is modulo \( N \).

An assertion \( J_{tok} \) says that the \( N \) most recent sends of token messages are totally ordered by causality. This is equivalent to stipulating that the timestamps on these token messages form an ascending sequence; for example, if \( tp \neq N - 1 \), then \( l_{x_{tp}} \leq l_{x_{tp-1}} \leq \cdots \leq l_{x_0} \leq l_{x_{N-1}} \leq l_{x_{N-2}} \leq \cdots \leq l_{x_{tp+1}} \). Formally,

\[
    J_{tok} \triangleq (\forall i \neq tp : l_{x_{i+1}} \leq l_{x_i})
\]

An assertion relating the timestamps of token messages to the timestamps of basic messages is also needed. For this, we use an assertion \( J_{bas} \), whose informal interpretation is as follows.

Let \( m \) be a basic message sent from \( i \) to \( k \) that was sent before the \( \alpha \)th transmission of the token by the sender. If \( m \) was sent in the same direction that the token travels (i.e., if \( k < i \)), then \( m \) must be delivered before the \( \alpha \)th transmission of the token by the receiver. If \( m \) was sent in the other direction (i.e., if \( i \leq k \)), then \( m \) must be delivered before the \( (\alpha + 1) \)th transmission of the token by the receiver. \( J_{bas} \) holds throughout execution of the algorithm because causally-ordered message-passing is used for all messages—the values of timestamps are consistent with this ordering. We formalize the assertion using an additional state function.

\( nlx_i \): The second largest timestamp in \( \{ m \in \bigcup_j a_j^{tok} \mid sender(m) = i \} \), if such a timestamp exists; otherwise 0.

\[
    J_{bas} \triangleq (\forall i, k : \forall m \in X_{i,k} : \\
    (k \leq tp < i \quad \Rightarrow \quad nlx_i < ts(m)) \\
    \land (k < i \land \neg(k \leq tp < i) \Rightarrow lx_i < ts(m)) \\
    \land (i \leq tp < k \quad \Rightarrow \quad lx_i < ts(m)) \\
    \land (i \leq k \land \neg(i \leq tp < k) \Rightarrow nlx_i < ts(m)))
\]

Assertions \( J_1 \), \( J_{bas} \), and \( J_{tok} \) contain all of the information about message-delivery order needed
for correct operation of the algorithm. We encapsulate this information as a single assertion \( J \):

\[
J \triangleq J_1 \land J_{bas} \land J_{tok}
\]

As with processes, a color, either black or white, is associated with the token. The color of the token is represented as before—black is encoded as true, and white is encoded as false. While in transit, this boolean value is included in each token message; while the token is visiting a process \( i \), a new variable \( t_i \) is used to store the color of the last token message received by process \( i \).

Given a boolean value \( c \), \( mtktok(c) \) denotes a token value whose color is \( c \). The color of the token is extracted using a selector \( tokval \). Thus, \( istok(mtktok(c)) = true \) and \( tokval(mtktok(c)) = c \). In each round, the token is initially white. It becomes black (if it isn’t already) when it visits a process \( i \) (i.e. \( h_i \) equals true) that is black (i.e. \( b_i \) equals true). Thus, the token becomes black when it visits a process that has sent a basic message since last sending a token message, and the current token color is given by:

\[
tc \triangleq \begin{cases} 
    ttp \lor btp & \text{if } h_{tp} \\
    tokval(data(m)) & \text{if } \neg h_{tp} \land m \in \sigma_{tp}^{tok} \land ts(m) = \mathcal{A}_{tp+1}
\end{cases}
\]

We also add to each process \( i \) a new variable \( y_i \), which is used for temporary storage of received values.

When the token returns to the initiator, if either the initiator or the token is black, then the initiator starts another round. If both are white, then the system is quiescent (i.e., \( Q \) holds).\(^8\) This fact is implied in the proof outlines of Figure 1 by the \( Q \) in the precondition for the second branch of the alternation statement \( RELAY_0 \).

The operation of the algorithm is succinctly characterized by \( K \), where \( K \triangleq K_1 \lor K_2 \lor K_3 \) and:

\[
K_1 \triangleq (\forall i > tp : q_i \land (\forall k : \chi_{i,k} = \emptyset)) \land (h_{tp} \Rightarrow (\forall k \geq tp : \chi_{tp,k} = \emptyset))
\]

\[
K_2 \triangleq (\exists i \leq tp : b_i)
\]

\[
K_3 \triangleq tc
\]

\(^8\)Here, the initiator does not take any special action when quiescence is detected. A round of communication could easily be added to notify each process that quiescence has been detected.
Informally, $K_1$ says that every process visited by the token in the current round is quiescent and no basic message sent by one of these processes is in transit. Moreover, if the token is visiting process $tp$, then no basic messages sent by process $tp$ are in transit to processes the token has visited in this round. $K_2$ says that some process not already visited by the token during the current round is black. Finally, $K_3$ says that the token is black.

Assertions $J$ and $K$ are not quite strong enough to prove correctness of the algorithm. An assertion $I$ that expresses several simple properties of the algorithm (e.g., that there is always at most one token message in the system) is also needed. Thus, we define $I \triangleq I \land J \land K$, where

$$
I \triangleq (\forall i:\quad (|\{i \mid \sigma_i^{tok} \neq \rho_i^{tok}\}| \leq 1)
\land (\forall m \in \sigma_i:\quad ts(m) \leq vt_{sender(m)})
\land (\forall m \in \sigma_i:\quad ts(m) \leq vt_i \implies m \in \rho_i)
\land (\forall m \in \rho_i:\quad ts(m) \leq vt_i)
\land (|\sigma_i^{tok} - \rho_i^{tok}| \leq 1)
\land ((h_i \lor \sigma_i^{tok} \neq \rho_i^{tok}) \implies tp = i)
\land (h_i \implies (\sigma_i^{tok} \neq \emptyset \land (\forall j : \sigma_j^{tok} = \rho_j^{tok})))
\land (\sigma_i^{tok} = \{m \in \bigcup_j \sigma_j^{tok} \mid sender(m) = i + 1\})
\land (total(\{m \in \bigcup_j \sigma_j \mid sender(m) = i\}))
\land (total(\bigcup_j \sigma_j^{tok}))
\land (lx_i \leq vt_i)
\land (x_{i,i} = \emptyset))
$$

and $total(S)$ holds iff $\{t \mid (\exists m \in S : ts(m) = t)\}$ is totally ordered by $\prec$.

Proof outlines for processes augmented to detect termination appear in Figure 1. The Appendix contains a detailed justification of the proof outlines.

Angle brackets indicate that the enclosed statement is executed atomically [Lam80].\footnote{Angle brackets are not actually necessary for correctness. They do simplify the proof slightly, so we have elected to use them.} Also, communication statements may appear in guards, so we use the following proof rule for iteration statements:
Proof Outline for Process $i$

$$\{I \land \neg h_i \land tp \geq i \land (i = 0 \Rightarrow (\forall j: \sigma_j^{tok} = \emptyset))\}$$

$\text{INIT}_i \quad \{I\}$

do

$\neg \exists j \neq i \quad \{I \land g_{ij}\}$

$g_{ij} \quad \{I \land g_{ij}\}$

$b_i := \text{true} \quad \{I \land g_j \land b_i\}$

send $e_{ij}$ to $j \quad \{I \land g_{ij}\}$

$S_{ij} \quad \{I\}$

receive $y_i \quad \{I \land (\neg istok(y_i) \Rightarrow K[q_i := \text{false}] \land (istok(y_i) \Rightarrow tp = i \land \neg h_i \land tc = tokval(y_i))\}$

if

$istok(y_i) \quad \{I \land tp = i \land \neg h_i \land tc = tokval(y_i)\}$

$h_i := \text{true} \quad \{I\}$

$t_i := tokval(y_i) \quad \{I\}$

fi

$q_i \land h_i \quad \{I \land q_i \land h_i\}$

RELAY$_i \quad \{I\}$

do

od

$\{I\}$

$\text{INIT}_0 \triangleq \text{send mktok(false) to } N - 1 \quad \{I \land \neg h_0 \land tp \geq 0\}$

$\text{Init}_0$

$\text{RELAY}_0 \triangleq \text{if } (t_0 \lor b_0) \quad \{I \land h_0\}$

$\langle\text{send mktok(false) to } N - 1 \land
\begin{align*}
h_0 & := \text{false} \\
\end{align*}
\rangle \quad \{I\}$

$b_0 := \text{false} \quad \{I\}$

$\neg(t_0 \lor b_0) \quad \{I \land Q\}$

(* quiescent *)

skip \quad \{I\}$

fi

For $0 < j < N$:

$\text{INIT}_j \triangleq \text{Init}_j$

$\text{RELAY}_j \triangleq \langle\text{send mktok(t}_j \lor b_j)\text{ to } j - 1 \land
\begin{align*}
h_j & := \text{false} \\
\end{align*}
\rangle

Figure 1: Proof Outlines
Iteration Rule:

For \(i \in [1..N]\), \(\{I \land g_i\} C_i \{P_i\} PO_i \{I\}\)

\[
\begin{align*}
\text{do} \\
\quad \forall i \in [1..N], \quad g_i; C_i \rightarrow \{P_i\} \\
\quad PO_i \\
\text{od} \\
\{I \land \neg(\forall i \in [1..N] \neg g_i)\}
\end{align*}
\]  

(14)

Here, \(g_i\) is a boolean expression and \(C_i\) is a receive or skip statement.\(^{10}\) One might expect there to be an assertion between \(g_i\) and \(C_i\) in the rule's conclusion. Expression \(g_i\) contains program variables of only process \(i\), so \(g_i\) cannot be invalidated by execution of another process. In particular, interference cannot occur even if evaluation of \(g_i\) and execution of \(C_i\) are not performed as a single indivisible action. Thus, there is no need to make the assertion explicit.

To illustrate reasoning about receive statements, we give a detailed proof for the triple

\[
\{I\} \text{receive} \ y_i \{I \land (\neg istok(y_i) \Rightarrow K[q_i := false]) \land (istok(y_i) \Rightarrow tp = i \land \neg h_i \land tc = tokval(y_i))\}
\]

(15)

This triple arises as a hypothesis in the application of the Iteration Rule to the main loop of each process. The triple expresses a crucial fact about the algorithm—that activation of a process (i.e., the changing of \(q_i\) to false) by reception of a basic message does not falsify \(K\). By Receive Rule (9), we can deduce (15) from

\[
I \land m_i \in \sigma_i - \rho_i \Rightarrow (I \land (\neg istok(y_i) \Rightarrow K[q_i := false]) \land (istok(y_i) \Rightarrow tp = i \land \neg h_i \land tc = tokval(y_i)))'
\]

(16)

where for any term \(t\),

\[
t' \triangleq t[y_i := data(m_i), \ \nu t_i := \max(\nu t_i, ts(m_i)), \ \rho_i := \rho_i \oplus m_i]
\]

We show in the Appendix that \(I \Rightarrow I'\) is valid. Here, we first show that

\[
I \land m_i \in \sigma_i - \rho_i \land \neg istok(y_i) \Rightarrow K[q_i := false]'
\]

(17)

\(^{10}\)The guard "\(g_i; \text{skip}\)" is abbreviated "\(g\)"; the guard "true; receive \(x\)" is abbreviated "receive \(x\)".
is valid. We assume the antecedent and prove the consequent. Note that

\[ K[q_i := false]' = (K'_1[q_i := false] \lor K_2 \lor K_3) \]

Thus, if \( K_2 \) or \( K_3 \) holds, then so does (17). Suppose neither \( K_2 \) nor \( K_3 \) holds. Since \( I \) holds by assumption, \( K \) must also hold, so \( K_1 \) must hold as well. We now show that in this case, \( K_1'[q_i := false] \) holds. First, note that \( K_1' \) holds; this follows easily from the fact that \( K_1 \) holds. The proof proceeds by case analysis on the relative values of \( i \) and \( tp \).

**case** \( i \leq tp \): \( K_1' \) does not depend on the \( q_j \)'s for \( j \leq tp \). Therefore, since \( K_1' \) holds, so does \( K_1'[q_i := false] \).

**case** \( i > tp \): We show that this case is impossible. Let \( k \triangleq \text{sender}(m_i) \). From the antecedent of (17) and the definition of \( \chi_{k,i} \), we conclude \( m_i \in \chi_{k,i} \).

**case** \( k \leq tp \): Instantiating the universally quantified variables \( i \) and \( k \) in \( J_{ba,k} \) with \( k \) and \( i \), respectively, we conclude (using the third conjunct of \( J_{ba,k} \)) that \( lz_k < ts(m_i) \). Using \( J_1 \), this implies that \( b_k \) holds, which implies that \( K_2 \) holds. This contradicts the assumption above that neither \( K_2 \) nor \( K_3 \) hold.

**case** \( k > tp \): By assumption, \( K_1 \) holds, so \((\forall j : \chi_{k,j} = \emptyset)\), so \( \chi_{k,i} = \emptyset \). From the antecedent of (17), we have \( \neg istok(y_i') \) (i.e., \( \neg istok(data(m_i)) \)) and \( m_i \in \sigma_i - \rho_i \), so by definition of \( \chi_{k,i} \), we have \( m_i \in \chi_{k,i} \), a contradiction.

Finally, consider showing that \((istok(y_i) \Rightarrow tp = i \land \neg h_i \land tc = tokval(y_i))' \) holds whenever the antecedent of (16) holds. This is equivalent to showing

\[ I \land m_i \in \sigma_i - \rho_i \land istok(data(m_i)) \Rightarrow tp = i \land \neg h_i \land tc = tokval(data(m_i)) \]

(18)

We assume the antecedent and prove the consequent. From the antecedent, we conclude \( m_i \in \sigma_i^{tok} - \rho_i^{tok} \). Thus, \( \sigma_i^{tok} \neq \rho_i^{tok} \), so by conjunct \((\forall i : (h_i \lor \sigma_i^{tok} \neq \rho_i^{tok}) \Rightarrow tp = i)\) in \( I \), \( tp = i \) holds. We next show, by contradiction, that \( \neg h_i \) holds. Suppose not; then \( h_i \) holds, so (using \( I \)), \( \sigma_i^{tok} = \rho_i^{tok} \), which contradicts \( m_i \in \sigma_i^{tok} - \rho_i^{tok} \). Finally, we show that \( tc = tokval(data(m_i)) \). From \( I \), \( |\sigma_i^{tok} - \rho_i^{tok}| \leq 1 \); thus, \( m_i \) is the only unreceived message in \( \sigma_i^{tok} \), so \( m_i \) must have the largest timestamp in \( \sigma_i^{tok} \), so \( ts(m_i) = lz_i+1 \). This, together with \( \neg h_i \), implies \( tc = tokval(data(m_i)) \).
Comparison to Related Work

The first correctness argument applicable to this distributed termination detection algorithm in an asynchronous setting is (to the best of our knowledge) an operational argument due to Raynal and Helary [RH90]. Proposition 3.8.1 in [RH90] establishes partial correctness assuming that the message-delivery order satisfies a property $P$. Our proof assumes causally-ordered delivery, which implies our predicate $J_{bas}$; $J_{bas}$ is similar to but slightly stronger than property $P$ of [RH90].

Another operational (albeit more formal) proof, by Charron-Bost et al., appears in [CBMT92]. It shows correctness of this termination detection algorithm for systems that communicate using causally-ordered message-passing. The proofs there differ considerably from the invariant-based analysis of the synchronous case in [DFvG83]. In fact, Charron-Bost et al. claim that correctness proofs for all algorithms that use causally-ordered delivery “must consider the execution as a whole, rather than concentrate on assertions that remain invariant in each global state” ([CBMT92], p. 34). The existence of our proof, which is an invariant-based analysis, refutes this claim.

5 Conclusions

We have presented a Hoare-style proof system for causally-ordered delivery. Through an example, we have demonstrated the feasibility of our approach to reasoning about causally-ordered delivery. The example, a distributed termination detection algorithm, has been treated using other approaches, so there is now an opportunity to compare those approaches with the one in this paper.

The fact that a correctness proof for causally-ordered delivery can be based closely on the analysis of a synchronous version is a significant benefit of the approach discussed in this paper. We support a two-step approach to verifying algorithms that use asynchronous message-passing [Gri90]:

1. Verify a synchronous version of the algorithm (presumably a simpler task).

2. Modify the algorithm and the proof to obtain a correctness proof for the asynchronous version of the algorithm.

One benefit of this two-step approach is that it leads naturally to a focus on and accurate determination of the ordering requirements needed by the algorithm. An interesting question is the extent to which this approach can be made formal and systematic.
References


A Proof of Correctness

We show that the proof outlines in Figure 1 are valid. We discuss only the triples for non-composite statements. It is easy to prove validity of the proof outlines in Figure 1 using these results and the inference rules for sequential composition, iteration, and alternation. The triples for non-composite statements that arise in the proofs for each process in isolation are listed in Figure 2. Proving invariance of $I$ is straightforward, so we omit those details. For brevity, we sometimes content ourselves with giving an informal explanation for why a triple is valid; based on this, the reader should have little difficulty constructing a formal proof.
For $0 < i < N$:

T1: $\{ I \land \neg h_i \land t_p \geq i \} \ Init_i \ {I}$

T2: $\{ I \land g_{ij} \} b_i := \text{true} \ \{ I \land g_{ij} \land b_i \}$

T3: $\{ I \land g_{ij} \land b_i \} \text{send} \ e_{ij} \ \text{to} \ j \ \{ I \land g_{ij} \}$

T4: $\{ I \land g_{ij} \} S_{ij} \ \{ I \}$

T5: $\{ I \} \text{receive} \ y_i \ \{ I \land (\neg istok(y_i) \Rightarrow K[q_i := \text{false}] \land (istok(y_i) \Rightarrow t_p = i \land \neg h_i \land tc = tokval(y_i)))\}$

T6: $\{ I \land t_p = i \land \neg h_i \land tc = tokval(y_i) \} \{ h_i := \text{true} \land t_i := tokval(y_i) \} \ {I}$

T7: $\{ I \land K[q_i := \text{false}] \} \ x_i := y_i \ \{ I \land K[q_i := \text{false}] \}$

T8: $\{ I \land K[q_i := \text{false}] \} \ R_i \ \{ I \}$

T9: $\{ I \land g_i \land h_i \} \{ \text{send} \ mktok(t_i \lor b_i) \ \text{to} \ i - 1 \land h_i := \text{false} \land b_i := \text{false} \} \ {I}$

T10: $\{ I \land \neg h_0 \land t_p \geq 0 \land (\forall j: \sigma_j^0 = \emptyset) \} \text{send} \ mktok(\text{false}) \ \text{to} \ N - 1 \ \{ I \land \neg h_0 \land t_p \geq 0 \}$

T11: $\{ I \land \neg h_0 \land t_p \geq 0 \} \ Init_0 \ \{ I \}$

T12: $\{ I \land h_0 \} \{ \text{send} \ mktok(\text{false}) \ \text{to} \ N - 1 \land h_0 := \text{false} \land b_0 := \text{false} \} \ {I}$

Figure 2: Triples for non-composite statements.

A.1 Proof for Process $i > 0$ in Isolation

T1: $\{ I \land \neg h_i \land t_p \geq i \} \ INIT_i \ {I}$

Since $i > 0$, $\ INIT_i$ is $\ Init_i$. $J$ is unaffected by execution of $\ Init_i$ because $\ Init_i$ neither sends nor receives messages. To see that $K$ is also unaffected, note that the only variables that appear in $K$ and can be assigned by $\ Init_i$ are those appearing in $q_i$, and that $K$ is independent of $q_i$ for $i \leq t_p$.

The precondition of T1 implies $i \leq t_p$, so $K$ is not invalidated by $\ INIT_i$.

T2: $\{ I \land g_{ij} \} \ b_i := \text{true} \ \{ I \land g_{ij} \land b_i \}$

$J$ is unaffected by execution of this statement. Variable $b_i$ occurs only positively in $K$, so setting $b_i$ to true never falsifies $K$. Finally, $b_i$ does not appear in $g_{ij}$, so the assignment to $b_i$ does not falsify $g_{ij}$.

T3: $\{ I \land g_{ij} \land b_i \} \text{send} \ e_{ij} \ \text{to} \ j \ \{ I \land g_{ij} \}$

We prove invariance of $J$ as follows. $J_1$ is preserved because $b_i$ holds. $J_{tok}$ is unaffected because the message being sent is not a token message. Let $m$ denote the element added to $\sigma_j$ by executing
this statement. To show that $J_{ba}$ is preserved, it suffices to show that $l x_i \prec ts(m)$ and $n l x_i \prec ts(m)$ hold, since $J_{ba}$ is then satisfied regardless of which conjunct applies to this message. By definition of the send statement, $ts(m) = inc(vt_i, i)$, so (by definition of $\prec$) $vt_i \prec ts(m)$. From $I$, we have $l x_i \preceq vt_i$, so by transitivity of $\prec$, $l x_i \prec ts(m)$. It follows from the definitions of $l x_i$ and $n l x_i$ that $n l x_i \preceq l x_i$, so by transitivity of $\prec$, $n l x_i \prec ts(m)$. Thus, $J_{ba}$ is preserved.

The proof that $K$ is preserved is by case analysis on the disjunct of $K$ that holds initially.

case $K_1$: In this case, $tp \geq i$ must also hold, since $i > tp$ and $K_1$ imply $q_i$, contradicting $g_{ij}$ in the precondition of T3. Since $i \leq tp$ and $b_i$ hold, $K_2$ also holds, so see that case.

case $K_2$: $K_2$ is unaffected by execution of this statement, so $K_2$ still holds after execution of this statement.

case $K_3$: $K_3$ is unaffected by execution of this statement, so $K_3$ still holds after execution of this statement.

T4 : \{I \land g_{ij}\} S_{ij} \{I\}

$J$ is unaffected by execution of $S_{ij}$ because $S_{ij}$ neither sends nor receives messages. The only variables that appear in $K$ and can be assigned by $S_{ij}$ are those appearing in $q_i$. Since $g_{ij}$ holds, $q_i$ is false, so execution of $S_{ij}$ either truthifies $q_i$ or leaves it unchanged. Variable $q_i$ occurs only positively in $K$, so truthifying $q_i$ never falsifies $K$.

T5 : \{I\} receive $y_i$ \{I \land (\neg istok(y_i) \Rightarrow K[q_i := false]) \land (istok(y_i) \Rightarrow tp = i \land \neg h_i \land tc = tokval(y_i))\}

Adding elements to $\rho_i$ never falsifies $J$ or $K$, and $J$ and $K$ do not depend on $y_i$ or $vt_i$, so $J$ and $K$ are preserved by execution of this statement. We argued in Section 4 that the other conjuncts in the postcondition hold after execution of this statement.

T6 : \{I \land tp = i \land \neg h_i \land tc = tokval(y_i)\} \{h_i := true \quad t_i := tokval(y_i)\} \{I\}

$J$ is unaffected by execution of this statement because messages are neither sent nor received. The proof that $K$ is preserved is by case analysis on the disjunct of $K$ that holds initially. Note that the only variables or state functions appearing in $K$ that are affected by execution of this
case $K_1$: The first conjunct of $K_1$ is unaffected by execution of this statement. We now consider the second conjunct. If $(\forall k \geq i : \chi_{i,k} = \emptyset)$, then, since $tp = i$ appears in the precondition, we can conclude that $K_1$ holds after $h_i$ is set to true by this statement. If $(\forall k \geq i : \chi_{i,k} = \emptyset)$ does not hold, then there exist $k$ and $m$ such that $k \geq i$ and $m \in \chi_{i,k}$. $I$ implies $\chi_{i,i} = \emptyset$, so it must be that $k > i$ and $m \in \chi_{i,k}$. From the precondition of this triple, $i = tp$, so $i \leq tp < k$. Thus, by the third conjunct of $J_{bas}$, $lx_i < ts(m)$, so by $J_1$, $b_i$ holds. Since $tp = i$ and $b_i$ hold, $K_2$ must hold, so see that case.

case $K_2$: $K_2$ is unaffected by execution of this statement, so $K_2$ still holds after execution of this statement.

case $K_3$: In this case, $tc$ holds. Let $m$ be the element of $\sigma_i^{tok}$ such that $ts(m) = lx_{i+1}$. Execution of this statement changes $tc$ from $tokval(y_i)$ to $tokval(y_i) \lor b_i$, so $K_3$ is not falsified.

$T7 : \{ I \land K[q_i := false] \} \ x_i := y_i \{ I \land K[q_i := false] \}$

$J$ is unaffected by execution of this statement because messages are neither sent nor received. Note that $x_i$ can appear in $K$ only in $q_i$. Since $K[q_i := false]$ holds before execution, and since $q_i$ occurs only positively in $K$, changing $q_i$ can’t falsify $K$. Finally, $K[q_i := false]$ is unaffected by execution of this statement.

$T8 : \{ I \land K[q_i := false] \} \ R_i \{ I \}$

$J$ is unaffected by execution of this statement because messages are neither sent nor received. The only variables that appear in $K$ and can be assigned by $R_i$ are those appearing in $q_i$. Since $q_i$ occurs only positively in $K$, and since $K$ holds even if $q_i$ doesn’t (because $K[q_i := false]$ appears in the precondition), execution of this statement cannot falsify $K$.

$T9 : \{ I \land q_i \land h_i \} \langle \text{send mktok}(t_i \lor b_i) \to i - 1 \quad h_i := false \quad b_i := false \rangle \{ I \}$

First, we show that execution of this statement changes $tp$ from $i$ to $i - 1$. Since $h_i$ holds, we conclude (using $I$) that $tp = i$. It follows from the definition of $tp$ that $(\forall j \neq i + 1 : lx_j < lx_{i+1})$.  

22
Since \( h_i \) holds, \( I \) implies \( \sigma_{i}^{tok} \neq \emptyset \) and \( \sigma_{i}^{tok} = \rho_{i}^{tok} \). Let \( m \) be the element of \( \sigma_{i}^{tok} \) with the largest timestamp; thus, \( z_{i+1} = ts(m) \). Since \( \sigma_{i}^{tok} = \rho_{i}^{tok}, m \in \rho_{i} \), so (using \( I \)) \( ts(m) \leq vt_i \), i.e., \( z_{i+1} \leq vt_i \). Thus, by transitivity of \( \prec \), \( (\forall j \neq i + 1: z_{j} \prec vt_i) \). Since this statement does not affect \( z_{j} \) for \( j \neq i \), after execution of this statement, \( (\forall j \not\in \{i, i + 1\}: z_{j} \prec vt_i) \) holds. After execution of this statement, \( z_{i} = inc(vt_i, i) \). By definition of \( \prec \), \( vt_i \prec inc(vt_i, i) \), so by transitivity, \( (\forall j \not\in \{i, i + 1\}: z_{j} \prec vt_i) \) holds after execution. Since \( z_{i+1} \leq vt_i \prec inc(vt_i, i) \), after execution, \( z_{i+1} \prec vt_i \) holds. Thus, after execution, \( (\forall j \neq i: z_{j} \prec z_{i}) \) holds, so by definition of \( tp, tp = i - 1 \).

\( J_1 \) is preserved because after execution of this statement, \( z_{i} \) is larger than the timestamps of all messages previously sent by process \( i \). To show that \( J_{tok} \) is preserved, it suffices to show \( z_{i+1} \leq inc(vt_i, i) \), since \( z_{i} = inc(vt_i, i) \) after execution. Let \( m \) be the member of \( \sigma_{i}^{tok} \) with the largest timestamp (this is well-defined since \( h_i \) and \( I \) imply that \( \sigma_{i}^{tok} \neq \emptyset \) and that the timestamps of messages in \( \sigma_{i}^{tok} \) are totally-ordered by \( \prec \)); thus, \( z_{i+1} = ts(m) \). Since \( h_i \) holds, we conclude using \( I \) that \( \sigma_{i}^{tok} = \rho_{i}^{tok} \), so \( m \in \rho_{i} \), which implies (using \( I \)) that \( ts(m) \leq vt_i \). By definition of \( \prec \), \( vt_i \prec inc(vt_i, i) \). Thus, \( z_{i+1} \leq vt_i \prec inc(vt_i, i) \).

Next we show that \( J_{bas} \) is preserved. Fix \( j, k \), and \( m \in \chi_{j,k} \) (we have renamed the bound variable \( i \) in (13) to \( j \)). We do a case analysis on the relative values of \( j, k, \) and \( tp \).

**Case** \( k \leq tp < j \): Since \( J_{bas} \) holds, \( n z_{j} \prec ts(m) \). If \( tp \neq k \), then \( k \leq tp < j \) is preserved by execution of this statement, so we must show \( n z_{j} \prec ts(m) \), which we already know to be true. Suppose \( tp = k \). After execution of this statement, \( \neg(k \leq tp < j) \), so we must show \( z_{j} \prec ts(m) \). We give a proof by contradiction: we suppose \( \neg(z_{j} \prec ts(m)) \) and show \( m \in \rho_{k} \), which contradicts the assumption \( m \in \chi_{j,k} \). \( J_{tok} \) implies \( z_{j} \leq z_{j-1} \cdots \leq z_{i+1} \), so \( ts(m) \leq z_{i+1} \). Let \( m' \) be the member of \( \sigma_{i}^{tok} \) with the largest timestamp (this is well-defined since \( h_i \) and \( I \) imply that \( \sigma_{i}^{tok} \neq \emptyset \) and that the timestamps of messages in \( \sigma_{i}^{tok} \) are totally-ordered by \( \prec \)); thus, \( z_{i+1} = ts(m') \), so \( ts(m) \leq ts(m') \). Since \( h_i \) holds, we conclude (using \( I \)) \( \sigma_{i}^{tok} = \rho_{i}^{tok} \), so (using \( I \)) \( m' \in \rho_{i} \), hence (again using \( I \)) \( ts(m') \leq vt_i \). Thus, \( ts(m) \leq ts(m') \leq vt_i \), so (using \( I \)) \( m \in \rho_{i} \). Since by assumption \( i = k, m \in \rho_{k} \).

**Case** \( k < j \) and \( \neg(k \leq tp < j) \): Since \( J_{bas} \) holds, \( z_{j} \prec ts(m) \). As in the previous case, preservation of \( J_{bas} \) is trivial if \( tp \neq j \). Suppose \( tp = j \). After execution of this statement, \( k \leq tp < j \), so we must show that \( n z_{j} \prec ts(m) \) then holds; this follows immediately from \( z_{j} \prec ts(m) \) and
the fact that the value of $n|x_j$ after execution of this statement equals the value of $|x_j$ before execution of this statement.

**case** $j \leq tp < k$: This case is analogous to the previous case.

**case** $j \leq k$ and $\neg(j \leq tp < k)$: This case is analogous to the first case.

Finally, we show that $K$ is preserved by execution of this statement. Recall that execution of this statement changes $tp$ from $i$ to $i - 1$. Note that execution of this statement leaves $tc$ unchanged. The proof that $K$ is preserved is by case analysis on the disjunct of $K$ that holds initially.

**case** $K_1$: We distinguish two subcases.

**case** $(\forall k : \chi_{i,k} = \emptyset)$: From the precondition of this triple, $q_i$ holds. Since execution of this statement does not affect $q_i$ or $\chi_{i,k}$ for all $k$, $K_1$ continues to hold after execution of this statement.

**case** $(\exists k : \chi_{i,k} \neq \emptyset)$: Since $K_1$ and $h_{tp}$ hold, $(\forall k \geq i : \chi_{i,k} = \emptyset)$ does too. This, together with the assumption $(\exists k : \chi_{i,k} \neq \emptyset)$, implies there exists $k$ such that $k < i$ and $\chi_{i,k} \neq \emptyset$. Let $m$ be an element of $\chi_{i,k}$. Since $k < i$ and $tp = i$, $J_{bas}$ implies $lx_i < ts(m)$, from which we conclude using $J_1$ that $b_i$ holds. After execution of this statement, $tc$ equals $t_i \lor b_i$, so $K_3$ then holds.

**case** $K_2$: Since $i = tp$, $K_2 = (\exists k < i : b_k) \lor b_i$. If the left disjunct holds, then $K_2$ still holds after execution of this statement. If the right disjunct holds before execution, then so does $K_3$ (because $h_i$ holds and $tp = i$), so see that case.

**case** $K_3$: $tc$ is unchanged by execution of this statement, so $K_3$ still holds after execution of this statement.

### A.2 Proof for Process 0 in Isolation

The verification of process $i$ when $i = 0$ in isolation involves the following triples, in addition to those discussed above.

\[ T10 : \{I \land \neg h_0 \land tp \geq 0 \land (\forall j : \sigma_j^{tok} = \emptyset)\} \text{send} mktok(false) \text{to} N - 1 \{I \land \neg h_0 \land tp \geq 0\} \]

First, we show that after execution of this statement, $tp = N - 1$. The precondition implies $(\forall j : \{m \in \bigcup_k \sigma_k^{tok} | \text{sender}(m) = j\} = \emptyset)$; it follows from the definition of $lx_j$ that $lx_j = \emptyset$ for all
j. After execution of this statement, $lx_0 = inc(vt_0, 0)$. From the definition of $\prec$, $\bar{0} \prec inc(vt, 0)$ for all vector times $vt$. From the definition of $tp$, we conclude that after execution of this statement, $(\forall j \neq 0 : lx_j \prec lx_0)$ holds, hence $tp = N - 1$.

$J_1$ is preserved because after execution of this statement, $lx_0$ is larger than the timestamps of all messages previously sent by process 0. To show that $J_{tok}$ holds after execution of this statement, we need to show that $\bar{0} \leq \bar{0}$ and $\bar{0} \leq inc(vt_0, 0)$; both of these facts follow from the definition of $\prec$. To see that $J_{bas}$ holds after execution of this statement, note that $lx_j = \bar{0}$ and (by the same reasoning) $nlx_j = 0$ for $j \neq 0$. Thus, $J_{bas}$ holds trivially for $j \neq 0$. For $j = 0$, note that there is no process $k$ such that $k < 0$, and recall that after execution of this statement, $tp = N - 1$. Thus, the only non-vacuous conjunct in $J_{bas}$ is the bottom one. This conjunct holds because $nlx_0 = 0$.

The conjunct $tp \geq 0$ in the postcondition holds after execution because $tp$ then equals $N - 1$, as shown above. Finally, note that $\neg h_0$ is unaffected by execution of this statement.

T11 : \{$I \land \neg h_0 \land tp \geq 0$\} $Init_0 \{I\}$

Validity of this triple follows by the same reasoning as for triple T1.

T12 : \{$I \land h_0$\} $<$send mktok(false) to N - 1 $>$ $h_0 := false$ $b_0 := false$ \{I\}$

$J$ is preserved by the same reasoning as for triple T9. We now show that execution of this statement truthifies $K_1$. Since $h_0$ holds, we conclude (using I) that $tp = 0$ holds before execution of this statement, so $\neg h_{N-1}$, because otherwise, $I$ implies $tp = N - 1$, which contradicts $tp = 0$. By the same reasoning as for triple T9, after execution of this statement, $tp = N - 1$. Thus, $K_1$ holds vacuously after execution of this statement.

Finally, we discuss one proof obligation that arises when using the foregoing results to verify the proof outlines given in Figure 1. When proving the second branch of $RELAY_0$, the following subgoal arises:

$I \land q_0 \land h_0 \land \neg(t_0 \lor b_0) \Rightarrow Q$

We assume the antecedent and prove the consequent. First, we show that $K_1$ must hold, by showing that $K_2$ and $K_3$ do not. Since $h_0$ holds, we conclude (using I) that $tp = 0$. From $tp = 0$ and $\neg b_0$, we conclude that $K_2$ does not hold. From $h_0$ and $\neg(t_0 \lor b_0)$, we conclude that $K_3$ does not hold. Thus, assuming the antecedent holds, $K_1$ also holds. It is easy to show that $K_1$ and the antecedent
together imply \( Q \).

A.3 Interference Freedom

Most of the interference freedom obligations can be discharged easily, using derived rules such as Interference Freedom for Synchronously Altered Assertions [LG81]. One non-trivial triple that arises in the proof of interference freedom is

\[
\{ K[q_j := \text{false}] \land K \land K[q_i := \text{false}] \} R_i \{ K[q_j := \text{false}] \}
\]

where \( j \neq i \). By the Assignment Axiom, validity of this triple follows from

\[
K[q_j := \text{false}] \land K \land K[q_i := \text{false}] \Rightarrow K[q_i := \text{false}, q_j := \text{false}]
\]

We assume the antecedent and prove the consequent. If \( K_2 \) holds, then \( K_2[q_i := \text{false}, q_j := \text{false}] \) holds, since \( q_i \) and \( q_j \) do not appear in \( K_2 \). The same reasoning applies to \( K_3 \). If neither \( K_2 \) nor \( K_3 \) hold, then \( K_1[q_j := \text{false}] \land K_1 \land K_1[q_i := \text{false}] \) must hold. We show by contradiction that this implies \( i \leq tp \). Suppose \( i > tp \); then

\[
K_1 = q_i \land (\forall k : \chi_{i,k} = \emptyset)
\]

\[
\land (\forall i' > tp : i' \neq i \Rightarrow q'_i \land (\forall k : \chi_{i',k} = \emptyset))
\]

\[
\land (h_{tp} \Rightarrow (\forall k \geq tp : \chi_{tp,k} = \emptyset))
\]

so \( K_1[q_i := \text{false}] = \text{false} \land \cdots \), so \( K_1[q_i := \text{false}] \) does not hold, which contradicts the assumption above. Thus, \( i \leq tp \). Analogous reasoning shows that \( j \leq tp \). Since \( i \leq tp \) and \( j \leq tp \), \( K_1 \) is independent of \( q_i \) and \( q_j \). By assumption, \( K_1 \) holds, so \( K_1[q_i := \text{false}, q_j := \text{false}] \) also holds.