Specifying & Refining a File Transfer Protocol

MEB: Chapter 4

EECS3342 Z: System Specification and Refinement
Winter 2022

CHEN-WEI WANG

Learning Outcomes

This module is designed to help you review:

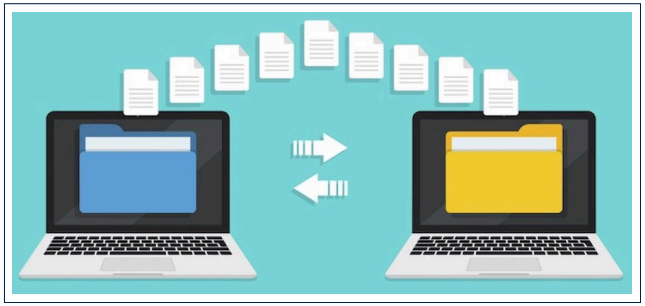
- What a ***Requirement Document*** (***RD***) is
- What a ***refinement*** is
- Writing ***formal specifications***
  - (Static) contexts: constants, axioms, theorems
  - (Dynamic) machines: variables, invariants, events, guards, actions
- ***Proof Obligations*** (***POs***) associated with proving:
  - ***refinements***
  - system ***properties***
- Applying ***inference rules*** of the ***sequent calculus***

A Different Application Domain

- The bridge controller we ***specified***, ***refined***, and ***proved*** exemplifies a ***reactive system***, working with the physical world via:
  - ***sensors*** [`a, b, c, ml_pass, il_pass`]
  - ***actuators*** [`ml_tl, il_tl`]
- We now study an example exemplifying a ***distributed program***:
  - A ***protocol*** followed by two ***agents***, residing on **distinct** geographical locations, on a computer **network**
  - Each file is transmitted ***asynchronously***: bytes of the file do **not** arrive at the ***receiver*** all at one go.
  - Language of ***predicates***, ***sets***, and ***relations*** required
  - The **same** principles of generating ***proof obligations*** apply.

Requirements Document: File Transfer Protocol (FTP)

You are required to implement a system for transmitting files between ***agents*** over a computer network.



Page Source: `https://www.venafi.com`

Requirements Document: R-Descriptions

Each ***R-Description*** is an **atomic** ***specification*** of an intended ***functionality*** or a desired ***property*** of the working system.

REQ1	The protocol ensures the copy of a file from the sender to the receiver.
REQ2	The file is supposed to be made of a sequence of items.
REQ3	The file is sent piece by piece between the two sites.

Refinement Strategy

- Recall the ***design strategy of progressive refinements***.
  0. ***initial model*** (m_0): a file is transmitted from the ***sender*** to the ***receiver***. [**REQ1**]
     However, at this ***most abstract*** model:
     - file transmitted from ***sender*** to ***receiver*** **synchronously** & **instantaneously**
     - transmission process ***abstracted*** away
  1. ***1st refinement*** (m_1 ***refining*** m_0):
     transmission is done ***asynchronously*** [**REQ2**, **REQ3**]
     However, at this more concrete model:
     - **no** communication between ***sender*** and ***receiver***
     - exchanges of ***messages*** and ***acknowledgements abstracted*** away
  2. ***2nd refinement*** (m_2 ***refining*** m_1):
     communication mechanism elaborated [**REQ2**, **REQ3**]
  3. ***final, 3rd refinement*** (m_3 ***refining*** m_2):
     communication mechanism optimized [**REQ2**, **REQ3**]
- Recall *Correct by Construction* :

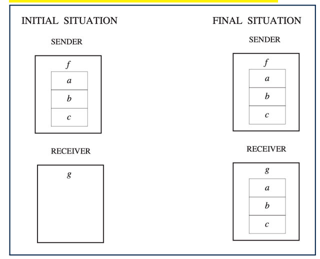
  From each ***model*** to its ***refinement***, only a manageable amount of details are added, making it ***feasible*** to conduct **analysis** and **proofs**.

Model m_0: Abstraction

- In this most ***abstract*** perception of the protocol, we do **not** consider the ***sender*** and ***receiver***:
  - residing in geographically distinct locations
  - communicating via message exchanges
- Instead, we focus on this single ***requirement***:

REQ1	The protocol ensures the copy of a file from the sender to the receiver.

- ***Abstraction Strategy*** :



  - Observe the system with the ***process of transmission abstracted*** away
  - **only** meant to inform ***what*** the protocol is supposed to achieve
  - **not** meant to detail ***how*** the transmission is achieved

Math Background Review

Refer to LECTURE 1 for reviewing:
- Predicates [e.g., ∀]
- Sets
- Relations and Operations
- Functions

Model m_0: Abstract State Space

1. The **static** part formulates the ***file*** (from the ***sender***'s end) as a sequence of data items:

sets: $D, BOOLEAN$	constants: n, f	axioms:
		axm0_1 : $n > 0$
		axm0_2 : $f \in 1..n \to D$
		axm0_3 : $BOOLEAN = \{TRUE, FALSE\}$

2. The **dynamic** part of the state consists of two ***variables***:

variables: g, b	invariants:
	inv0_1a : $g \in 1..n \rightarrowtail D$
	inv0_1b : $b \in BOOLEAN$
	inv0_2 : ??
	inv0_3 : ??

- ✓ ***g***: file from the ***receiver***'s end
- ✓ ***b***: whether or not the ***transmission*** is completed
- ✓ **inv0_1a** and **inv0_1b** are ***typing*** constraints.
- ✓ **inv0_2** specifies what happens **before** the transmission
- ✓ **inv0_3** specifies what happens **after** the transmission

Model m_0: State Transitions via Events

- The system acts as an **ABSTRACT STATE MACHINE (ASM)**: it ***evolves*** as ***actions of enabled events*** change values of variables, subject to ***invariants***.
- Initially, before the transmission:

```
init
  begin
    ??
  end
```

  - Nothing has been transmitted to the ***receiver***.
  - The ***transmission*** process has not been completed.
- Finally, after the transmission:

```
final
  when
    ??
  then
    ??
  end
```

  - The entire file f has been transmitted to the ***receiver***.
  - The ***transmission*** process has been completed.
  - In this ***abstract*** model:
    - Think of the transmission being ***instantaneous***.
    - A later ***refinement*** specifies how f is transmitted ***asynchronously***.

PO of Invariant Establishment

- How many ***sequents*** to be proved? [# invariants]
- We have four ***sequents*** generated for ***event*** *init* of model m_0:

1. $n > 0$, $f \in 1..n \to D$, $BOOLEAN = \{TRUE, FALSE\}$ $\vdash$ $\varnothing \in 1..n \rightarrowtail D$ — **init/inv0_1a/INV**
2. $n > 0$, $f \in 1..n \to D$, $BOOLEAN = \{TRUE, FALSE\}$ $\vdash$ $FALSE \in BOOLEAN$ — **init/inv0_1b/INV**
3. $n > 0$, $f \in 1..n \to D$, $BOOLEAN = \{TRUE, FALSE\}$ $\vdash$ $FALSE = FALSE \Rightarrow \varnothing = \varnothing$ — **init/inv0_2/INV**
4. $n > 0$, $f \in 1..n \to D$, $BOOLEAN = \{TRUE, FALSE\}$ $\vdash$ $FALSE = TRUE \Rightarrow \varnothing = f$ — **init/inv0_3/INV**

- Exercises: Prove the above sequents related to ***invariant establishment***.

PO of Invariant Preservation

- How many ***sequents*** to be proved? [# non-init events × # invariants]
- We have four ***sequents*** generated for ***event*** *final* of model m_0:

final/inv0_1a/INV: $n > 0$, $f \in 1..n \to D$, $BOOLEAN = \{TRUE, FALSE\}$, $g \in 1..n \rightarrowtail D$, $b \in BOOLEAN$, $b = FALSE \Rightarrow g = \varnothing$, $b = TRUE \Rightarrow g = f$, $b = FALSE$ $\vdash$ $f \in 1..n \rightarrowtail D$

final/inv0_1b/INV: $n > 0$, $f \in 1..n \to D$, $BOOLEAN = \{TRUE, FALSE\}$, $g \in 1..n \rightarrowtail D$, $b \in BOOLEAN$, $b = FALSE \Rightarrow g = \varnothing$, $b = TRUE \Rightarrow g = f$, $b = FALSE$ $\vdash$ $TRUE \in BOOLEAN$

final/inv0_2/INV: $n > 0$, $f \in 1..n \to D$, $BOOLEAN = \{TRUE, FALSE\}$, $g \in 1..n \rightarrowtail D$, $b \in BOOLEAN$, $b = FALSE \Rightarrow g = \varnothing$, $b = TRUE \Rightarrow g = f$, $b = FALSE$ $\vdash$ $TRUE = FALSE \Rightarrow f = \varnothing$

final/inv0_3/INV: $n > 0$, $f \in 1..n \to D$, $BOOLEAN = \{TRUE, FALSE\}$, $g \in 1..n \rightarrowtail D$, $b \in BOOLEAN$, $b = FALSE \Rightarrow g = \varnothing$, $b = TRUE \Rightarrow g = f$, $b = FALSE$ $\vdash$ $TRUE = TRUE \Rightarrow f = f$

- Exercises: Prove the above sequents related to ***invariant preservation***.

Initial Model: Summary

- Our ***initial model*** m_0 is ***provably correct*** w.r.t.:
  - Establishment of ***Invariants***
  - Preservation of ***Invariants***
  - ***Deadlock*** Freedom [EXERCISE]
- Here is the ***specification*** of m_0:

sets: $D, BOOLEAN$

constants: n, f

axioms:
- **axm0_1** : $n > 0$
- **axm0_2** : $f \in 1 .. n \rightarrow D$
- **axm0_3** : $BOOLEAN = \{TRUE, FALSE\}$

variables: g, b

invariants:
- **inv0_1a** : $g \in 1 .. n \mapsto D$
- **inv0_1b** : $b \in BOOLEAN$
- **inv0_2** : $b = FALSE \Rightarrow g = \varnothing$
- **inv0_3** : $b = TRUE \Rightarrow g = f$

init
begin
$g := \varnothing$
$b := FALSE$
end

final
when
$b = FALSE$
then
$g := f$
$b := TRUE$
end

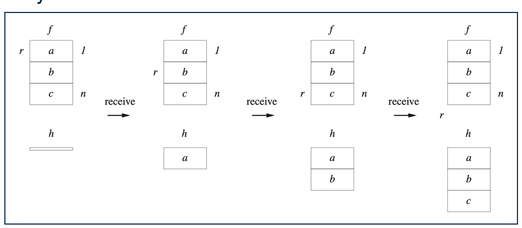
Model m_1: "More Concrete" Abstraction

- In $\mathbf{m_0}$, the transmission (evt. `final`) is ***synchronous*** and ***instantaneous***.
- The 1st ***refinement*** has a more ***concrete*** perception of the file transmission:
  - The sender's file is coped gradually, ***element by element***, to the receiver.
    → Such progress is denoted by occurrences of a ***new event*** `receive`.

h: elements transmitted so far

r: index of element to be sent

abstract variable ***g*** is replaced by ***concrete*** variables ***h*** and ***r***.



  - Nonetheless, communication between two agents remain ***abstracted*** away!
- That is, we focus on these two ***intended functionalities***:

REQ2	The file is supposed to be made of a sequence of items.
REQ3	The file is sent piece by piece between the two sites.

- We are ***obliged to prove*** this ***added concreteness*** is ***consistent*** with m_0.

Model m_1: Refined, Concrete State Space

1. The **static** part remains the same as $\mathbf{m_0}$:

sets: $D, BOOLEAN$

constants: n, f

axioms:
- **axm0_1** : $n > 0$
- **axm0_2** : $f \in 1 .. n \rightarrow D$
- **axm0_3** : $BOOLEAN = \{TRUE, FALSE\}$

2. The **dynamic** part formulates the ***gradual*** transmission process:

variables: b, h, r

invariants:
- **inv1_1** : $r \in 1 .. n+1$
- **inv1_2** : ??
- **inv1_3** : ??
- **thm1_1** : ??

- **inv1_1**: typing constraint
- **inv2_2**: elements up to index ***r - 1*** have been transmitted
- **inv2_3**: transmission completed **means** no more elements to be transmitted
- **thm1_1**: transmission completed **means** receiver has a complete copy of sender's file
- A ***theorem***, once proved as ***derivable from invariants***, needs **not** be proved for ***preservation*** by events.

Model m_1: Property Provable from Invariants

- To prove that a ***theorem*** can be derived from the ***invariants***:

variables: b, h, r

invariants:
- ***inv1_1*** : $r \in 1 .. n+1$
- ***inv1_2*** : $h = (1 .. r-1) \lhd f$
- ***inv1_3*** : $b = TRUE \Rightarrow r = n+1$
- ***thm1_1*** : $b = TRUE \Rightarrow h = f$

- We need to prove the following ***sequent***:

$$r \in 1 .. n+1$$
$$h = (1 .. r-1) \lhd f$$
$$b = TRUE \Rightarrow r = n+1$$
$$\vdash$$
$$b = TRUE \Rightarrow h = f$$

- Exercise: Prove the above sequent.

Model m_1: Old and New Concrete Events

- Initially, before the transmission:

```
init
  begin
    ??
  end
```

  - The ***transmission*** process has not been completed.
  - Nothing has been transmitted to the ***receiver***.
  - First file element is available for transmission.
- While the transmission is ongoing:

```
receive
  when
    ??
  then
    ??
  end
```

  - **While** sender has **more** file elements available for transmission:
    - Next file element is received and ***accumulated*** to the receiver's copy.
    - Sender's ***next available*** file element is updated.
  - In this ***concrete*** model:
    - Receiver having access to sender's private variable r is ***unrealistic***.
    - A later ***refinement*** specifies how two agents communicate.
- Finally, after the transmission:

```
final
  when
    ??
  then
    ??
  end
```

  - **When** sender has **no** more file element available for transmission:
    - The ***transmission*** process is marked as completed.

PO of Invariant Establishment

- How many ***sequents*** to be proved? [# invariants]
- We have three ***sequents*** generated for ***event*** *init* of model m_1:

1. **init/inv1_1/INV**

$$n > 0 \\ f \in 1 \,..\, n \rightarrow D \\ BOOLEAN = \{TRUE, FALSE\} \\ \vdash \\ 1 \in 1 \,..\, n+1$$

2. **init/inv1_2/INV**

$$n > 0 \\ f \in 1 \,..\, n \rightarrow D \\ BOOLEAN = \{TRUE, FALSE\} \\ \vdash \\ \varnothing \in (1 \,..\, 1-1) \lhd f$$

3. **init/inv1_3/INV**

$$n > 0 \\ f \in 1 \,..\, n \rightarrow D \\ BOOLEAN = \{TRUE, FALSE\} \\ \vdash \\ FALSE = TRUE \Rightarrow 1 = n+1$$

- Exercises: Prove the above sequents related to ***invariant establishment***.

PO of Invariant Preservation – `final`

- We have three ***sequents*** generated for ***old event*** *final* of model m_1.
- Here is one of the sequents:

final/inv1_1/INV

$$n > 0 \\ f \in 1 \,..\, n \rightarrow D \\ BOOLEAN = \{TRUE, FALSE\} \\ g \in 1 \,..\, n \nrightarrow D \\ b \in BOOLEAN \\ b = FALSE \Rightarrow g = \varnothing \\ b = TRUE \Rightarrow g = f \\ r \in 1 \,..\, n+1 \\ h = (1 \,..\, r-1) \lhd f \\ b = TRUE \Rightarrow r = n+1 \\ b = FALSE \\ r = n+1 \\ \vdash \\ r \in 1 \,..\, n+1$$

- Exercises: Formulate & prove other sequents of ***invariant preservation***.

PO of Invariant Preservation – `receive`

- We have three ***sequents*** generated for ***new event*** *receive* of model m_1:

receive/inv1_1/INV

$$n > 0 \\ f \in 1 \,..\, n \rightarrow D \\ BOOLEAN = \{TRUE, FALSE\} \\ g \in 1 \,..\, n \nrightarrow D \\ b \in BOOLEAN \\ b = FALSE \Rightarrow g = \varnothing \\ b = TRUE \Rightarrow g = f \\ r \in 1 \,..\, n+1 \\ h = (1 \,..\, r-1) \lhd f \\ b = TRUE \Rightarrow r = n+1 \\ r \leq n \\ \vdash \\ (r+1) \in 1 \,..\, n+1$$

receive/inv1_2/INV

$$n > 0 \\ f \in 1 \,..\, n \rightarrow D \\ BOOLEAN = \{TRUE, FALSE\} \\ g \in 1 \,..\, n \nrightarrow D \\ b \in BOOLEAN \\ b = FALSE \Rightarrow g = \varnothing \\ b = TRUE \Rightarrow g = f \\ r \in 1 \,..\, n+1 \\ h = (1 \,..\, r-1) \lhd f \\ b = TRUE \Rightarrow r = n+1 \\ r \leq n \\ \vdash \\ h \cup \{(r, f(r))\} = (1 \,..\, (r+1)-1) \lhd f$$

receive/inv1_3/INV

$$n > 0 \\ f \in 1 \,..\, n \rightarrow D \\ BOOLEAN = \{TRUE, FALSE\} \\ g \in 1 \,..\, n \nrightarrow D \\ b \in BOOLEAN \\ b = FALSE \Rightarrow g = \varnothing \\ b = TRUE \Rightarrow g = f \\ r \in 1 \,..\, n+1 \\ h = (1 \,..\, r-1) \lhd f \\ b = TRUE \Rightarrow r = n+1 \\ r \leq n \\ \vdash \\ b = TRUE \Rightarrow (r+1) = n+1$$

- Exercises: Prove the above sequents of ***invariant preservation***.

Proving Refinement: receive/inv1_1/INV

$$\begin{array}{l} n > 0 \\ f \in 1 \,..\, n \rightarrow D \\ BOOLEAN = \{TRUE, FALSE\} \\ g \in 1 \,..\, n \nrightarrow D \\ b \in BOOLEAN \\ b = FALSE \Rightarrow g = \varnothing \\ b = TRUE \Rightarrow g = f \\ r \in 1 \,..\, n+1 \\ h = (1 \,..\, r-1) \lhd f \\ b = TRUE \Rightarrow r = n+1 \\ r \leq n \\ \vdash \\ (r+1) \in 1 \,..\, n+1 \end{array}$$

MON

$$\begin{array}{l} r \in 1 \,..\, n+1 \\ r \leq n \\ \vdash \\ (r+1) \in 1 \,..\, n+1 \end{array}$$

ARI

$$\begin{array}{l} 1 \leq r \wedge r \leq n+1 \\ r \leq n \\ \vdash \\ \quad 1 \leq (r+1) \\ \wedge \;\; (r+1) \leq n+1 \end{array}$$

AND_L

$$\begin{array}{l} 1 \leq r \\ r \leq n+1 \\ r \leq n \\ \vdash \\ \quad 1 \leq (r+1) \\ \wedge \;\; (r+1) \leq n+1 \end{array}$$

AND_R

$$\begin{array}{l} 1 \leq r \\ r \leq n+1 \\ r \leq n \\ \vdash \\ 1 \leq (r+1) \end{array}$$

MON

$$\begin{array}{l} 1 \leq r \\ \vdash \\ 1 \leq (r+1) \end{array}$$

ARI

$$\begin{array}{l} 1 \leq r \\ r \leq n+1 \\ r \leq n \\ \vdash \\ (r+1) \leq n+1 \end{array}$$

MON

$$\begin{array}{l} r \leq n \\ \vdash \\ (r+1) \leq n+1 \end{array}$$

ARI

$$\begin{array}{l} r \leq n \\ \vdash \\ r \leq n \end{array}$$

HYP

Proving Refinement: receive/inv1_2/INV

$$\begin{array}{l} n > 0 \\ f \in 1 \,..\, n \rightarrow D \\ BOOLEAN = \{TRUE, FALSE\} \\ g \in 1 \,..\, n \nrightarrow D \\ b \in BOOLEAN \\ b = FALSE \Rightarrow g = \varnothing \\ b = TRUE \Rightarrow g = f \\ r \in 1 \,..\, n+1 \\ h = (1 \,..\, r-1) \lhd f \\ b = TRUE \Rightarrow r = n+1 \\ r \leq n \\ \vdash \\ h \cup \{(r, f(r))\} = (1 \,..\, (r+1)-1) \lhd f \end{array}$$

MON

$$\begin{array}{l} f \in 1 \,..\, n \rightarrow D \\ r \in 1 \,..\, n+1 \\ h = (1 \,..\, r-1) \lhd f \\ r \leq n \\ \vdash \\ h \cup \{(r, f(r))\} = (1 \,..\, (r+1)-1) \lhd f \end{array}$$

ARI

$$\begin{array}{l} f \in 1 \,..\, n \rightarrow D \\ 1 \leq r \\ h = (1 \,..\, r-1) \lhd f \\ r \leq n \\ \vdash \\ h \cup \{(r, f(r))\} = (1 \,..\, (r+1)-1) \lhd f \end{array}$$

EQ_LR, MON, ARI

$$\begin{array}{l} f \in 1 \,..\, n \rightarrow D \\ 1 \leq r \\ r \leq n \\ \vdash \\ (1 \,..\, r-1) \lhd f \cup \{(r, f(r))\} = (1 \,..\, r) \lhd f \end{array}$$

ARI

Proving Refinement: receive/inv1_3/INV

$$\begin{array}{l} n > 0 \\ f \in 1 \,..\, n \rightarrow D \\ BOOLEAN = \{TRUE, FALSE\} \\ g \in 1 \,..\, n \nrightarrow D \\ b \in BOOLEAN \\ b = FALSE \Rightarrow g = \varnothing \\ b = TRUE \Rightarrow g = f \\ r \in 1 \,..\, n+1 \\ h = (1 \,..\, r-1) \lhd f \\ b = TRUE \Rightarrow r = n+1 \\ r \leq n \\ \vdash \\ b = TRUE \Rightarrow (r+1) = n+1 \end{array}$$

MON

$$\begin{array}{l} b = TRUE \Rightarrow r = n+1 \\ r \leq n \\ \vdash \\ b = TRUE \Rightarrow (r+1) = n+1 \end{array}$$

IMP_R

$$\begin{array}{l} b = TRUE \Rightarrow r = n+1 \\ r \leq n \\ b = TRUE \\ \vdash \\ (r+1) = n+1 \end{array}$$

IMP_L

$$\begin{array}{l} r = n+1 \\ r \leq n \\ b = TRUE \\ \vdash \\ (r+1) = n+1 \end{array}$$

EQ_LR, MON

$$\begin{array}{l} n+1 \leq n \\ b = TRUE \\ \vdash \\ ((n+1)+1) = n+1 \end{array}$$

ARI, MON

$$\begin{array}{l} \bot \\ \vdash \\ ((n+1)+1) = n+1 \end{array}$$

FALSE_L

m_1: PO of Convergence of New Events

- Recall:
  - Interleaving of ***new*** events charactered as an integer expression: ***variant***.
  - A variant $V(c, w)$ may refer to constants and/or ***concrete*** variables.
  - For m_1, let's try **variants** : $n + 1 - r$
- Accordingly, for the ***new*** event *receive*:

$$\begin{array}{l} n > 0 \\ f \in 1 \,..\, n \rightarrow D \\ BOOLEAN = \{TRUE, FALSE\} \\ g \in 1 \,..\, n \nrightarrow D \\ b \in BOOLEAN \\ b = FALSE \Rightarrow g = \varnothing \\ b = TRUE \Rightarrow g = f \\ r \in 1 \,..\, n+1 \\ h = (1 \,..\, r-1) \lhd f \\ b = TRUE \Rightarrow r = n+1 \\ r \leq n \\ \vdash \\ n+1-(r+1) < n+1-r \end{array}$$

receive/VAR

Exercises: Prove **receive/VAR** and Formulate/Prove **receive/NAT**.

First Refinement: Summary

- The ***first refinement*** m_1 is ***provably correct*** w.r.t.:
  - Establishment of ***Concrete Invariants*** [*init*]
  - Preservation of ***Concrete Invariants*** [old & new events]
  - Strengthening of ***guards*** [old events, EXERCISE]
  - ***Convergence*** (a.k.a. livelock freedom, non-divergence) [new events, EXERCISE]
  - Relative ***Deadlock*** Freedom [EXERCISE]
- Here is the ***specification*** of m_1:

```
sets: D, BOOLEAN          constants: n, f

axioms:
  axm0_1 : n > 0
  axm0_2 : f ∈ 1 .. n → D
  axm0_3 : BOOLEAN = {TRUE, FALSE}

variables:     invariants:
  b, h, r        inv1_1 : r ∈ 1 .. n + 1
                 inv1_2 : h = (1 .. r − 1) ◁ f
                 inv1_3 : b = TRUE ⇒ r = n + 1
                 thm1_1 : b = TRUE ⇒ h = f

init
  begin
    b := FALSE
    h := ∅
    r := 1
  end

final
  when
    r = n + 1
    b = FALSE
  then
    b := TRUE
  end

receive
  when
    r ≤ n
  then
    h := h ∪ {(r, f(r))}
    r := r + 1
  end

variants:
  n + 1 − r
```


Index (1)

Index (2)

Index (3)