

# Specifying & Refining a File Transfer Protocol

MEB: Chapter 4



EECS3342 Z: System  
Specification and Refinement  
Winter 2023

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**

2 of 28

## A Different Application Domain



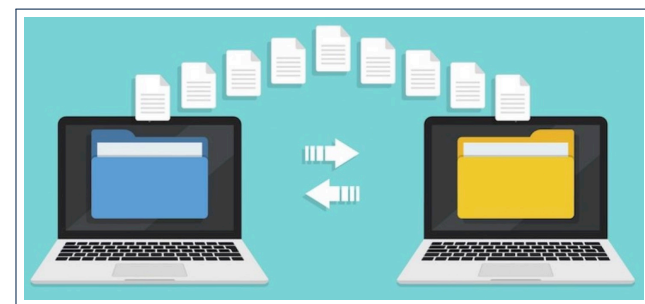
- The bridge controller we **specified**, **refined**, and **proved** exemplifies a **reactive system**, working with the physical world via:
  - **sensors** [ a, b, c, ml\_pass, il\_pass ]
  - **actuators** [ ml\_t1, il\_t1 ]
- We now study an example exemplifying a **distributed program** :
  - A **protocol** 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.

3 of 28

## 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>

4 of 28

# 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.

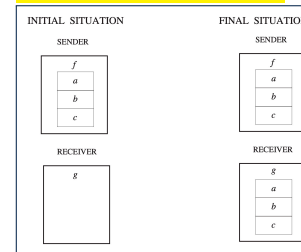
# Model $m_0$ : Abstraction



- 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

# Refinement Strategy



- Recall the **design strategy of progressive refinements**.
  - 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
  - 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
  - 2nd refinement** ( $m_2$  **refining**  $m_1$ ):  
communication mechanism **elaborated** [ REQ2, REQ3 ]
  - 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**.

# Math Background Review



Refer to LECTURE 1 for reviewing:

- Predicates
- Sets
- Relations and Operations
- Functions

[ e.g.,  $\forall$  ]

## Model $m_0$ : Abstract State Space

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

sets: $D, \text{BOOLEAN}$	constants: $n, f$	axioms: $\text{axm0.1} : n > 0$ $\text{axm0.2} : f \in 1..n \rightarrow D$ $\text{axm0.3} : \text{BOOLEAN} = \{ \text{TRUE}, \text{FALSE} \}$
---------------------------	-------------------	--

- The **dynamic** part of the state consists of two **variables**:

variables: $g, b$	invariants: $\text{inv0.1a} : g \in 1..n \rightarrow D$ $\text{inv0.1b} : b \in \text{BOOLEAN}$ $\text{inv0.2} : ??$ $\text{inv0.3} : ??$
-------------------	---

- ✓  $g$ : 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

9 of 28

## PO of Invariant Establishment

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

1.	$n > 0$ $f \in 1..n \rightarrow D$ $\text{BOOLEAN} = \{ \text{TRUE}, \text{FALSE} \}$ $\vdash$ $\emptyset \in 1..n \rightarrow D$	init/inv0.1a/INV
2.	$n > 0$ $f \in 1..n \rightarrow D$ $\text{BOOLEAN} = \{ \text{TRUE}, \text{FALSE} \}$ $\vdash$ $\text{FALSE} \in \text{BOOLEAN}$	init/inv0.1b/INV
3.	$n > 0$ $f \in 1..n \rightarrow D$ $\text{BOOLEAN} = \{ \text{TRUE}, \text{FALSE} \}$ $\vdash$ $\text{FALSE} = \text{FALSE} \Rightarrow \emptyset = \emptyset$	init/inv0.2/INV
4.	$n > 0$ $f \in 1..n \rightarrow D$ $\text{BOOLEAN} = \{ \text{TRUE}, \text{FALSE} \}$ $\vdash$ $\text{FALSE} = \text{TRUE} \Rightarrow \emptyset = f$	init/inv0.3/INV

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

11 of 28

## 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**.

10 of 28

## PO of Invariant Preservation

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

$n > 0$ $f \in 1..n \rightarrow D$ $\text{BOOLEAN} = \{ \text{TRUE}, \text{FALSE} \}$ $g \in 1..n \rightarrow D$ $b \in \text{BOOLEAN}$ $b = \text{FALSE} \Rightarrow g = \emptyset$ $b = \text{TRUE} \Rightarrow g = f$ $b = \text{FALSE}$ $\vdash$ $f \in 1..n \rightarrow D$	final/inv0.1a/INV	$n > 0$ $f \in 1..n \rightarrow D$ $\text{BOOLEAN} = \{ \text{TRUE}, \text{FALSE} \}$ $g \in 1..n \rightarrow D$ $b \in \text{BOOLEAN}$ $b = \text{FALSE} \Rightarrow g = \emptyset$ $b = \text{TRUE} \Rightarrow g = f$ $b = \text{FALSE}$ $\vdash$ $\text{TRUE} \in \text{BOOLEAN}$	final/inv0.1b/INV
$n > 0$ $f \in 1..n \rightarrow D$ $\text{BOOLEAN} = \{ \text{TRUE}, \text{FALSE} \}$ $g \in 1..n \rightarrow D$ $b \in \text{BOOLEAN}$ $b = \text{FALSE} \Rightarrow g = \emptyset$ $b = \text{TRUE} \Rightarrow g = f$ $b = \text{FALSE}$ $\vdash$ $\text{TRUE} = \text{FALSE} \Rightarrow f = \emptyset$	final/inv0.2/INV	$n > 0$ $f \in 1..n \rightarrow D$ $\text{BOOLEAN} = \{ \text{TRUE}, \text{FALSE} \}$ $g \in 1..n \rightarrow D$ $b \in \text{BOOLEAN}$ $b = \text{FALSE} \Rightarrow g = \emptyset$ $b = \text{TRUE} \Rightarrow g = f$ $b = \text{FALSE}$ $\vdash$ $\text{TRUE} = \text{TRUE} \Rightarrow f = f$	final/inv0.3/INV

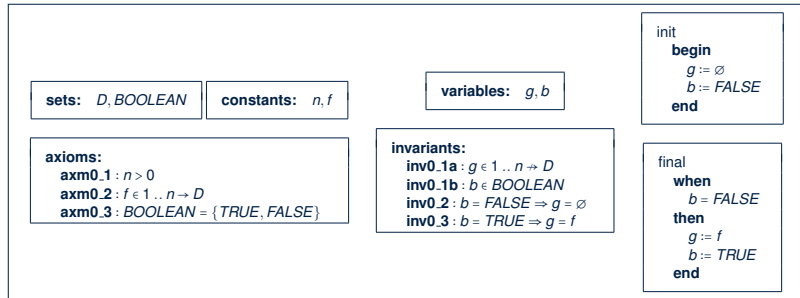
- Exercises:** Prove the above sequents related to **invariant preservation**.

12 of 28

## Initial Model: Summary

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

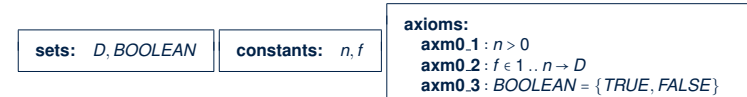
[ EXERCISE ]



13 of 28

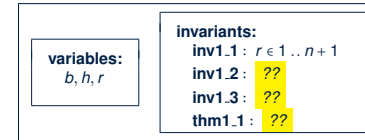
## Model $m_1$ : Refined, Concrete State Space

- The **static** part remains the same as  $m_0$ :



- The **dynamic** part formulates the *gradual* transmission process:

- $\text{inv1.1}$ : typing constraint
- $\text{inv1.2}$ : elements up to index  $r - 1$  have been transmitted
- $\text{inv1.3}$ : transmission completed **means** no more elements to be transmitted
- $\text{thm1.1}$ : 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.



15 of 28

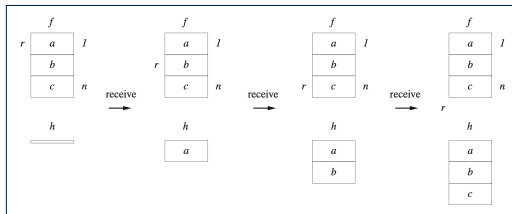
## Model $m_1$ : "More Concrete" Abstraction

- In  $m_0$ , 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 copied *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*:

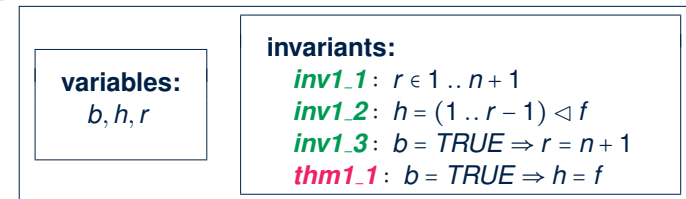
REQ2	The file is supposed to be made of a sequence of items.
REQ3	The file is sent piece by piece between the two sites.

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

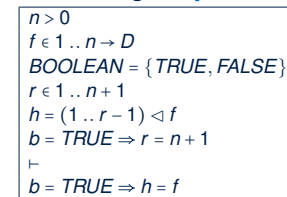
14 of 28

## Model $m_1$ : Property Provable from Invariants

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



- We need to prove the following *sequent*:



- Exercise**: Prove the above sequent.

16 of 28

## 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.

17 of 28

## PO of Invariant Preservation – final



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

```
n > 0
f ∈ 1 .. n → D
BOOLEAN = { TRUE, FALSE }
g ∈ 1 .. n → D
b ∈ BOOLEAN
b = FALSE ⇒ g = ∅
b = TRUE ⇒ g = f
r ∈ 1 .. n + 1
h = (1 .. r - 1) < f
b = TRUE ⇒ r = n + 1
b = FALSE
r = n + 1
⊢
r ∈ 1 .. n + 1
```

final/inv1\_1/INV

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

19 of 28

## PO of Invariant Establishment



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

1.  $n > 0$   
 $f \in 1 .. n \rightarrow D$   
 $BOOLEAN = \{ TRUE, FALSE \}$   
 $\vdash$   
 $1 \in 1 .. n + 1$       init/inv1\_1/INV

2.  $n > 0$   
 $f \in 1 .. n \rightarrow D$   
 $BOOLEAN = \{ TRUE, FALSE \}$   
 $\vdash$   
 $\emptyset \in (1 .. 1 - 1) < f$       init/inv1\_2/INV

3.  $n > 0$   
 $f \in 1 .. n \rightarrow D$   
 $BOOLEAN = \{ TRUE, FALSE \}$   
 $\vdash$   
 $FALSE = TRUE \Rightarrow 1 = n + 1$       init/inv1\_3/INV

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

18 of 28

## 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 ∈ 1 .. n → D
BOOLEAN = { TRUE, FALSE }
g ∈ 1 .. n → D
b ∈ BOOLEAN
b = FALSE ⇒ g = ∅
b = TRUE ⇒ g = f
r ∈ 1 .. n + 1
h = (1 .. r - 1) < f
b = TRUE ⇒ r = n + 1
r ≤ n
⊢
(r + 1) ∈ 1 .. n + 1
```

receive/inv1\_2/INV

```
n > 0
f ∈ 1 .. n → D
BOOLEAN = { TRUE, FALSE }
g ∈ 1 .. n → D
b ∈ BOOLEAN
b = FALSE ⇒ g = ∅
b = TRUE ⇒ g = f
r ∈ 1 .. n + 1
h = (1 .. r - 1) < f
b = TRUE ⇒ r = n + 1
r ≤ n
⊢
h ∪ { (r, f(r)) } = (1 .. (r + 1) - 1) < f
```

receive/inv1\_3/INV

```
n > 0
f ∈ 1 .. n → D
BOOLEAN = { TRUE, FALSE }
g ∈ 1 .. n → D
b ∈ BOOLEAN
b = FALSE ⇒ g = ∅
b = TRUE ⇒ g = f
r ∈ 1 .. n + 1
h = (1 .. r - 1) < f
b = TRUE ⇒ r = n + 1
r ≤ n
⊢
b = TRUE ⇒ (r + 1) = n + 1
```

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

20 of 28

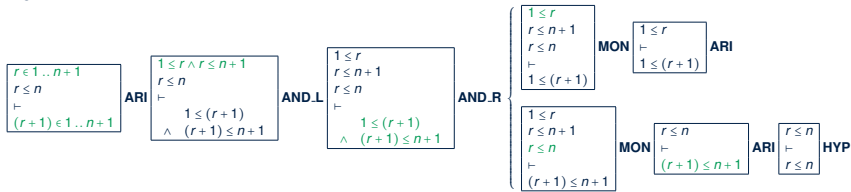
## Proving Refinement: receive/inv1\_1/INV



```

n > 0
f ∈ 1..n → D
BOOLEAN = { TRUE, FALSE }
g ∈ 1..n → D
b ∈ BOOLEAN
b = FALSE ⇒ g = ∅
b = TRUE ⇒ g = f
r ∈ 1..n+1
h = (1..r-1) < f
b = TRUE ⇒ r = n+1
r ≤ n
⊢
(r+1) ∈ 1..n+1
    
```

MON



21 of 28

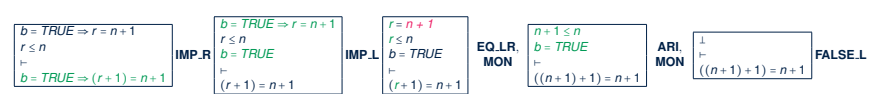
## Proving Refinement: receive/inv1\_3/INV



```

n > 0
f ∈ 1..n → D
BOOLEAN = { TRUE, FALSE }
g ∈ 1..n → D
b ∈ BOOLEAN
b = FALSE ⇒ g = ∅
b = TRUE ⇒ g = f
r ∈ 1..n+1
h = (1..r-1) < f
b = TRUE ⇒ r = n+1
r ≤ n
⊢
b = TRUE ⇒ (r+1) = n+1
    
```

MON



23 of 28

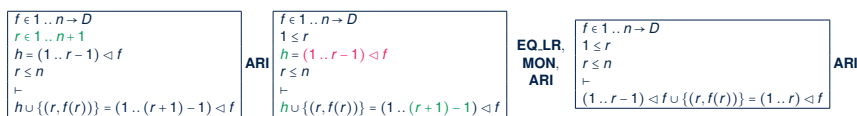
## Proving Refinement: receive/inv1\_2/INV



```

n > 0
f ∈ 1..n → D
BOOLEAN = { TRUE, FALSE }
g ∈ 1..n → D
b ∈ BOOLEAN
b = FALSE ⇒ g = ∅
b = TRUE ⇒ g = f
r ∈ 1..n+1
h = (1..r-1) < f
b = TRUE ⇒ r = n+1
r ≤ n
⊢
h ∪ {(r, f(r))} = (1..(r+1)-1) < f
    
```

MON



22 of 28

## $m_1$ : PO of Convergence of New Events



- Recall:
  - Interleaving of *new* events characterized 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*:

```

n > 0
f ∈ 1..n → D
BOOLEAN = { TRUE, FALSE }
g ∈ 1..n → D
b ∈ BOOLEAN
b = FALSE ⇒ g = ∅
b = TRUE ⇒ g = f
r ∈ 1..n+1
h = (1..r-1) < f
b = TRUE ⇒ r = n+1
r ≤ n
⊢
n+1 - (r+1) < n+1 - r
    
```

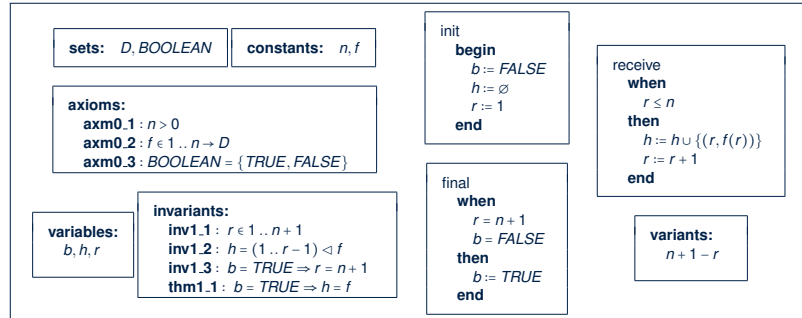
receive/VAR

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

24 of 28

## 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  $m_1$ :



25 of 28

## Index (2)

PO of Invariant Preservation

Initial Model: Summary

Model  $m_1$ : "More Concrete" Abstraction

Model  $m_1$ : Refined, Concrete State Space

Model  $m_1$ : Property Provable from Invariants

Model  $m_1$ : Old and New Concrete Events

PO of Invariant Establishment

PO of Invariant Preservation – final

PO of Invariant Preservation – receive

Proving Refinement: receive/inv1\_1/INV

Proving Refinement: receive/inv1\_2/INV

27 of 28

## Index (1)

Learning Outcomes

A Different Application Domain

Requirements Document:

File Transfer Protocol (FTP)

Requirements Document: R-Descriptions

Refinement Strategy

Model  $m_0$ : Abstraction

Math Background Review

Model  $m_0$ : Abstract State Space

Model  $m_0$ : State Transitions via Events

PO of Invariant Establishment

26 of 28

## Index (3)

Proving Refinement: receive/inv1\_3/INV

$m_1$ : PO of Convergence of New Events

First Refinement: Summary

28 of 28