

# Explicit Representation of the Heap

P. H. Schmitt, Benjamin Weiß, Mattias Ulbrich

Institut für Theoretische Informatik  
Fakultät für Informatik  
Universität Karlsruhe (TH)

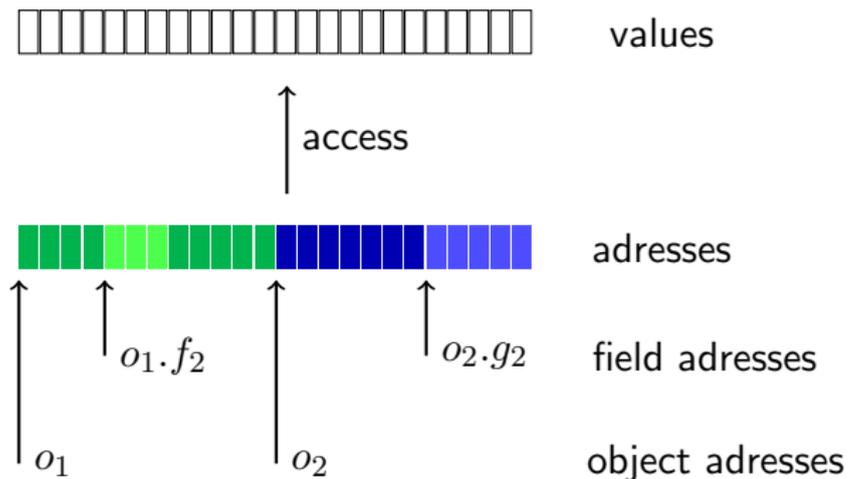


8. KeY-Symposium, May 19, 2009



# Views of the Heap

## Low Level View



# Views of the Heap

## Abstract View

$v_1$        $v_2$        $v_3$        $v_4$        $v_5$       values

↑  
access

$o_1.f_1$     $o_1.f_2$     $o_1.f_3$     $o_2.g_1$     $o_2.g_2$    locations



# Representations of the Heap

## Options

$D$  is the universe of all objects and values.

1. For every field  $f$  there is a function  $I(f) : D \rightarrow D$ .  
A heap state is the collection of all functions  $I(f)$ .  
Used by [ASM](#), [JACK](#), present [KeY](#).



# Representations of the Heap

## Options

$D$  is the universe of all objects and values.

1. For every field  $f$  there is a function  $I(f) : D \rightarrow D$ .  
A heap state is the collection of all functions  $I(f)$ .  
Used by [ASM](#), [JACK](#), present [KeY](#).
2. A heap state is represented by one partial function

$$\begin{aligned} H & : D \times Fields \rightarrow D \\ H & : Locations \rightarrow D \end{aligned}$$

used by [Boogie](#), [ESC2/Java](#), [Krakatoa](#), [KIV](#).



# Representations of the Heap

## Options

$D$  is the universe of all objects and values.

1. For every field  $f$  there is a function  $I(f) : D \rightarrow D$ .  
A heap state is the collection of all functions  $I(f)$ .  
Used by [ASM](#), [JACK](#), present [KeY](#).
2. A heap state is represented by one partial function

$$\begin{aligned} H & : D \times Fields \rightarrow D \\ H & : Locations \rightarrow D \end{aligned}$$

used by [Boogie](#), [ESC2/Java](#), [Krakatoa](#), [KIV](#).

3. Theoretically there are other options.



# New Proposal

## Changes to the Syntax

1. The only non-rigid symbols are constants corresponding to program variables.



# New Proposal

## Changes to the Syntax

1. The only non-rigid symbols are constants corresponding to program variables.
2. We consider only elementary updates with sequential and parallel composition.  
Maybe quantified and guarded updates are no longer necessary.



# New Proposal

## Changes to the Syntax

1. The only non-rigid symbols are constants corresponding to program variables.
2. We consider only elementary updates with sequential and parallel composition.  
Maybe quantified and guarded updates are no longer necessary.
3. There are two new types *Heap* and *Field* and an implicit program variable  $H$ .



# New Proposal

## Changes to the Syntax

1. The only non-rigid symbols are constants corresponding to program variables.
2. We consider only elementary updates with sequential and parallel composition.  
Maybe quantified and guarded updates are no longer necessary.
3. There are two new types *Heap* and *Field* and an implicit program variable  $H$ .
4. There are function symbols

$select : Heap \times Object \times Field \rightarrow Any$

$store : Heap \times (Object \times Field) \times Any \rightarrow Heap$



# New Proposal

## Changes to the Syntax

1. The only non-rigid symbols are constants corresponding to program variables.
2. We consider only elementary updates with sequential and parallel composition.  
Maybe quantified and guarded updates are no longer necessary.
3. There are two new types *Heap* and *Field* and an implicit program variable  $H$ .
4. There are function symbols

$select : Heap \times Object \times Field \rightarrow Any$

$store : Heap \times (Object \times Field) \times Any \rightarrow Heap$



# New Proposal

## Changes to the Syntax

1. The only non-rigid symbols are constants corresponding to program variables.
2. We consider only elementary updates with sequential and parallel composition.  
Maybe quantified and guarded updates are no longer necessary.
3. There are two new types *Heap* and *Field* and an implicit program variable  $H$ .
4. There are function symbols

$select : Heap \times Object \times Field \rightarrow Any$

$store : Heap \times (Object \times Field) \times Any \rightarrow Heap$

We ignore the problem of typing *select* for the moment.  
Maybe we can use Philip's idea for Boogie.



# New Proposal

## Semantics

- ▶ The states of a DL Kripke structure are given by the state vectors of the non-rigid constants (including  $\mathbb{H}$ ).



# New Proposal

## Semantics

- ▶ The states of a DL Kripke structure are given by the state vectors of the non-rigid constants (including  $\mathbb{H}$ ).
- ▶ Non-rigid constants take values in the *computational domain*  $\mathcal{D}$ .



# New Proposal

## Semantics

- ▶ The states of a DL Kripke structure are given by the state vectors of the non-rigid constants (including  $\mathbb{H}$ ).
- ▶ Non-rigid constants take values in the *computational domain*  $\mathcal{D}$ .
- ▶  $\mathcal{D}$  obviously satisfies the constant domain assumption.



# New Proposal

## Semantics

- ▶ The states of a DL Kripke structure are given by the state vectors of the non-rigid constants (including  $\mathbb{H}$ ).
- ▶ Non-rigid constants take values in the *computational domain*  $\mathcal{D}$ .
- ▶  $\mathcal{D}$  obviously satisfies the constant domain assumption.
- ▶ If we use underspecification of partial functions this amounts to considering different computational domains.



# New Proposal

## Axioms

- ▶ Different constants  $f, g$  of type *Field* have to be interpreted as different objects. Thus  $f \neq g$  is an axiom.



# New Proposal

## Axioms

- ▶ Different constants  $f, g$  of type *Field* have to be interpreted as different objects. Thus  $f \neq g$  is an axiom.
- ▶  $\forall \text{Heap } h; \forall \text{Object } o; \forall \text{Field } f; \forall \text{Any } x;$   
 $\text{select}(\text{store}(h, o, f, x), o, f) = x$



# New Proposal

## Axioms

- ▶ Different constants  $f, g$  of type *Field* have to be interpreted as different objects. Thus  $f \neq g$  is an axiom.
- ▶  $\forall Heap\ h; \forall Object\ o; \forall Field\ f; \forall Any\ x;$   
 $select(store(h, o, f, x), o, f) = x$



# New Proposal

## Axioms

- ▶ Different constants  $f, g$  of type *Field* have to be interpreted as different objects. Thus  $f \neq g$  is an axiom.
- ▶  $\forall \text{Heap } h; \forall \text{Object } o; \forall \text{Field } f; \forall \text{Any } x;$   
 $\text{select}(\text{store}(h, o, f, x), o, f) = x$

This is an instance of the theory of arrays, supported by many SMT solvers.



# New Proposal

## Axioms

- ▶ Different constants  $f, g$  of type *Field* have to be interpreted as different objects. Thus  $f \neq g$  is an axiom.
- ▶  $\forall \text{Heap } h; \forall \text{Object } o; \forall \text{Field } f; \forall \text{Any } x;$   
 $\text{select}(\text{store}(h, o, f, x), o, f) = x$

This is an instance of the theory of arrays, supported by many SMT solvers.

It consists of the sorts *Array*, *Index* and *Element*, of the two function symbols  $\text{select} : \text{Array} \times \text{Index} \rightarrow \text{Element}$  and  $\text{store} : \text{Array} \times \text{Index} \times \text{Element} \rightarrow \text{Array}$ , and the axiom:

$$\forall \text{Array } a; \forall \text{Index } i; \forall \text{Element } e; \forall \text{Index } j;$$
$$\text{select}(\text{store}(a, i, e), j) \doteq \text{if}(i \doteq j) \text{then}(e) \text{else}(\text{select}(a, j))$$



# New Proposal

## Pretty Printing

- ▶ We can render  $select(H, o, f)$  as  $o.f$ .



# New Proposal

## Pretty Printing

- ▶ We can render  $select(H, o, f)$  as  $o.f$ .
- ▶ An update  $H := store(H, o, f, x)$  can be rendered as  $o.f := x$ .



# New Proposal

## Pretty Printing

- ▶ We can render  $select(H, o, f)$  as  $o.f$ .
- ▶ An update  $H := store(H, o, f, x)$  can be rendered as  $o.f := x$ .
- ▶ As a generalisation of the above, an update

$$H := store(\dots (store(H, o_1, f_1, x_1), \dots), o_n, f_n, x_n)$$

can be rendered as a fake “parallel update”

$$o_1.f_1 := x_1 \parallel \dots \parallel o_n.f_n := x_n$$



# A Frist Example

## Present Version

`[o.f = 1; o.g = 2;](o.f ≐ 1)`



# A First Example

## Present Version

$[o.f = 1; o.g = 2;](o.f \doteq 1)$

↓ assignment

$\{o.f := 1\}[o.g = 2;](o.f \doteq 1)$

(1)



# A First Example

## Present Version

$$[o.f = 1; o.g = 2;](o.f \doteq 1) \quad (1)$$

$\downarrow$  assignment

$$\{o.f := 1\}[o.g = 2;](o.f \doteq 1) \quad (2)$$

$\downarrow$  assignment

$$\{o.f := 1\}\{o.g := 2\}[](o.f \doteq 1)$$



# A First Example

## Present Version

$$[o.f = 1; o.g = 2;](o.f \doteq 1) \quad (1)$$

↓ assignment

$$\{o.f := 1\}[o.g = 2;](o.f \doteq 1) \quad (2)$$

↓ assignment

$$\{o.f := 1\}\{o.g := 2\}[](o.f \doteq 1) \quad (3)$$

↓ update simplification

$$\{o.f := 1 \parallel o.g := 2\}[](o.f \doteq 1)$$



# A Frist Example

## Present Version

$$[o.f = 1; o.g = 2;](o.f \dot{=} 1) \quad (1)$$

↓ assignment

$$\{o.f := 1\}[o.g = 2;](o.f \dot{=} 1) \quad (2)$$

↓ assignment

$$\{o.f := 1\}\{o.g := 2\}[](o.f \dot{=} 1) \quad (3)$$

↓ update simplification

$$\{o.f := 1 \parallel o.g := 2\}[](o.f \dot{=} 1) \quad (4)$$

↓ empty modality, update simplification

$$1 \dot{=} 1$$



# A Frist Example

## Present Version

$$[o.f = 1; o.g = 2;](o.f \doteq 1) \quad (1)$$

↓ assignment

$$\{o.f := 1\}[o.g = 2;](o.f \doteq 1) \quad (2)$$

↓ assignment

$$\{o.f := 1\}\{o.g := 2\}[](o.f \doteq 1) \quad (3)$$

↓ update simplification

$$\{o.f := 1 \parallel o.g := 2\}[](o.f \doteq 1) \quad (4)$$

↓ empty modality, update simplification

$$1 \doteq 1 \quad (5)$$

↓ close

\*



## A Frist Example (New Version)

$[o.f = 1; o.g = 2;](o.f \doteq 1)$      $[o.f = 1; o.g = 2;](select(H, o, f) \doteq 1)$



## A Frist Example (New Version)

$[o.f = 1; o.g = 2;](o.f \doteq 1) \quad [o.f = 1; o.g = 2;](select(H, o, f) \doteq 1)$   
assignment  $\downarrow$   
 $\{o.f := 1\}[o.g = 2;](o.f \doteq 1)$   
 $\{H := store(H, o, f, 1)\}[o.g = 2;](select(H, o, f) \doteq 1)$



## A First Example (New Version)

$[o.f = 1; o.g = 2;](o.f \doteq 1) \quad [o.f = 1; o.g = 2;](select(H, o, f) \doteq 1)$   
assignment  $\downarrow$   
 $\{o.f := 1\}[o.g = 2;](o.f \doteq 1)$   
assignment  $\downarrow$   $\{H := store(H, o, f, 1)\}[o.g = 2;](select(H, o, f) \doteq 1)$   
 $\{o.f := 1\}\{o.g := 2\}[](o.f \doteq 1)$   
 $\{H := store(H, o, f, 1)\}\{H := store(H, o, g, 2)\}[](\dots)$

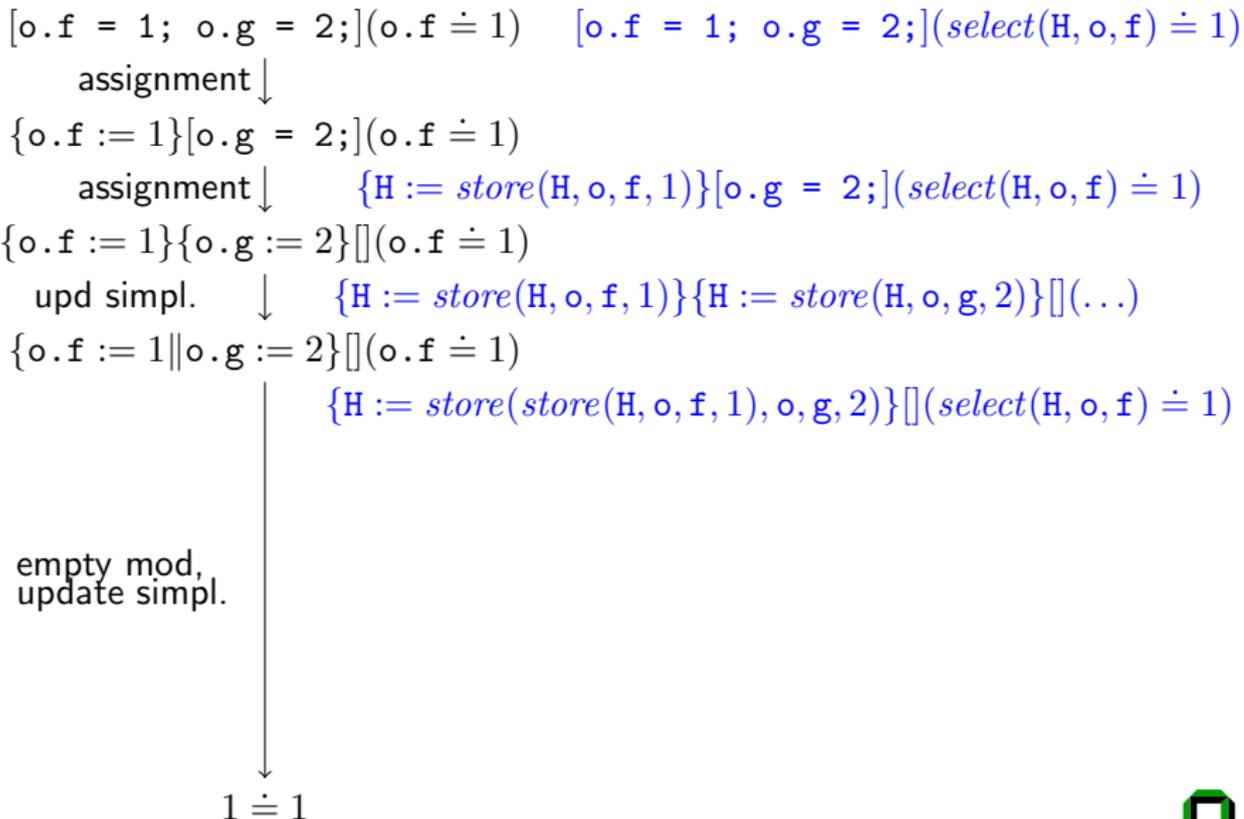


## A Frist Example (New Version)

$$\begin{array}{l} [o.f = 1; o.g = 2;](o.f \doteq 1) \quad [o.f = 1; o.g = 2;](select(H, o, f) \doteq 1) \\ \text{assignment} \downarrow \\ \{o.f := 1\}[o.g = 2;](o.f \doteq 1) \\ \text{assignment} \downarrow \quad \{H := store(H, o, f, 1)\}[o.g = 2;](select(H, o, f) \doteq 1) \\ \{o.f := 1\}\{o.g := 2\}[](o.f \doteq 1) \\ \text{upd simpl.} \quad \downarrow \quad \{H := store(H, o, f, 1)\}\{H := store(H, o, g, 2)\}[](\dots) \\ \{o.f := 1 || o.g := 2\}[](o.f \doteq 1) \\ \quad \quad \quad \{H := store(store(H, o, f, 1), o, g, 2)\}[](select(H, o, f) \doteq 1) \end{array}$$



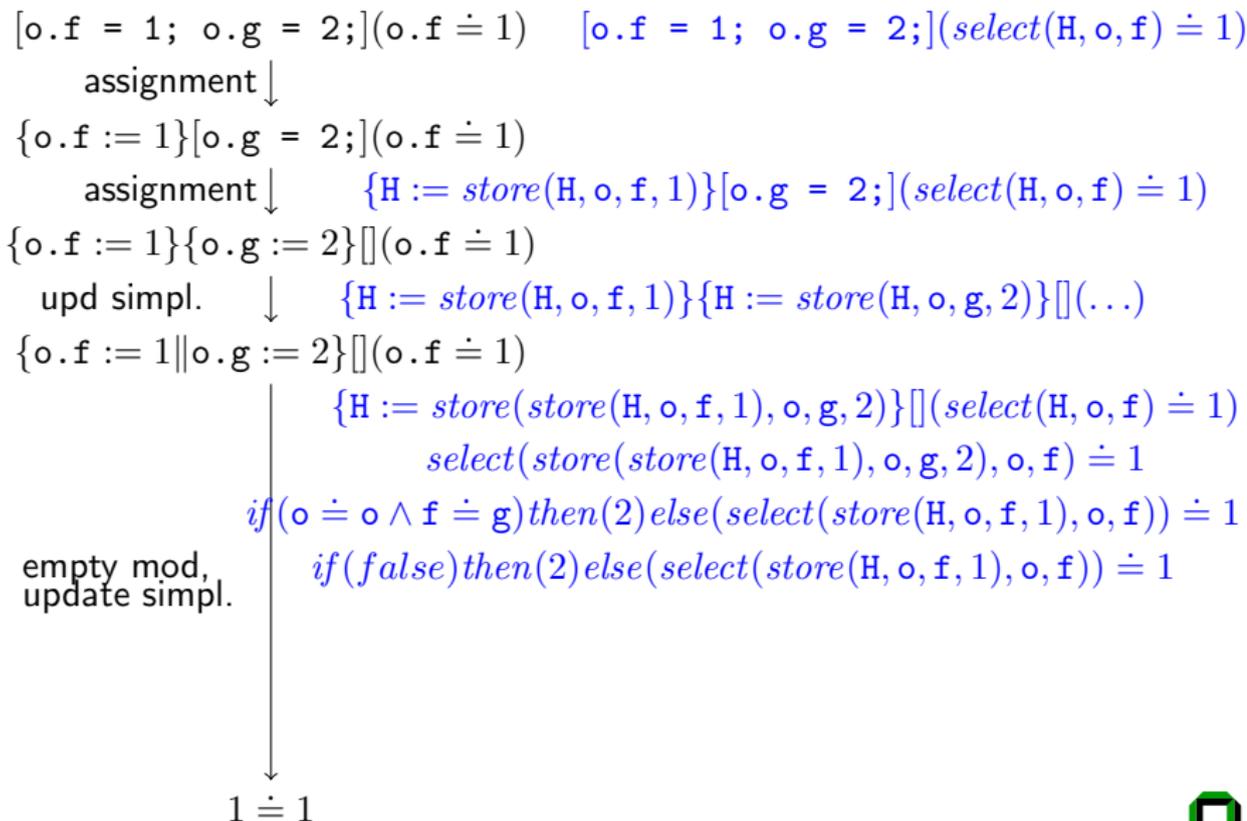
# A First Example (New Version)



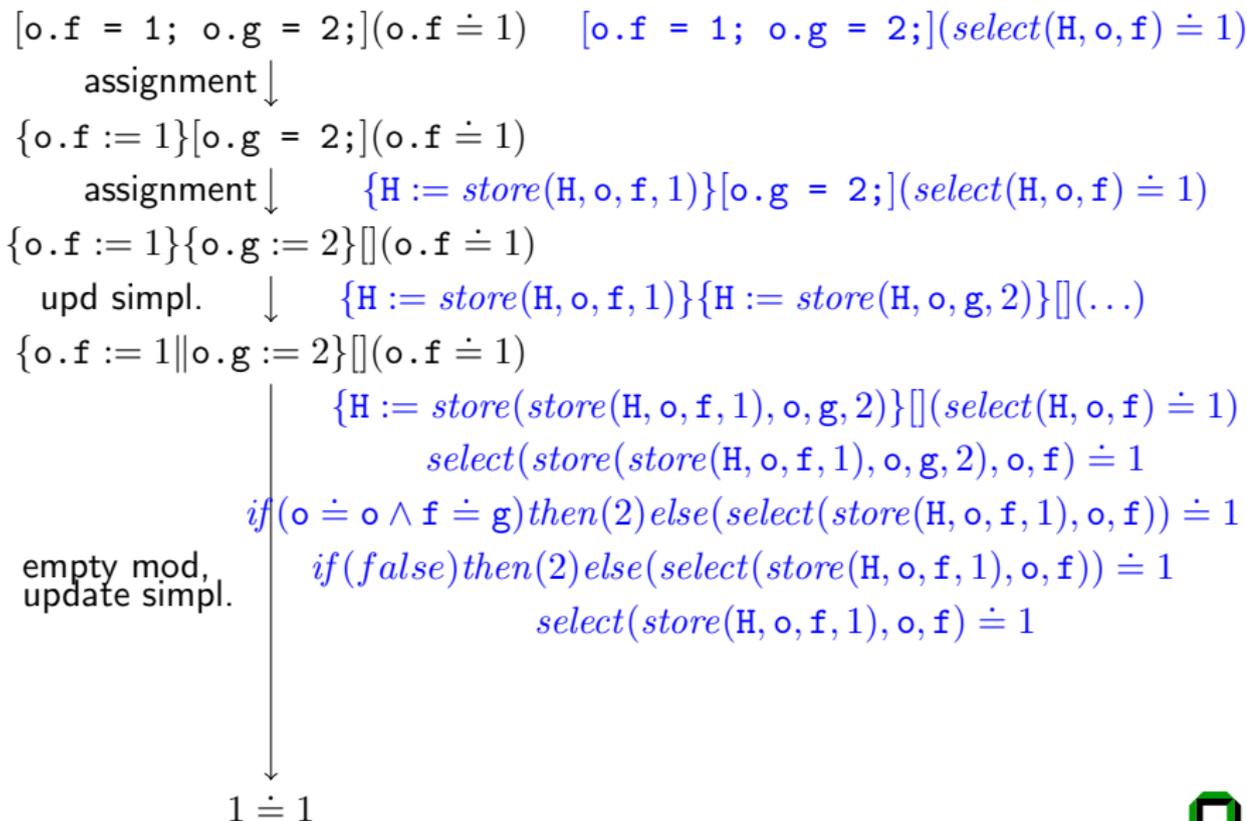




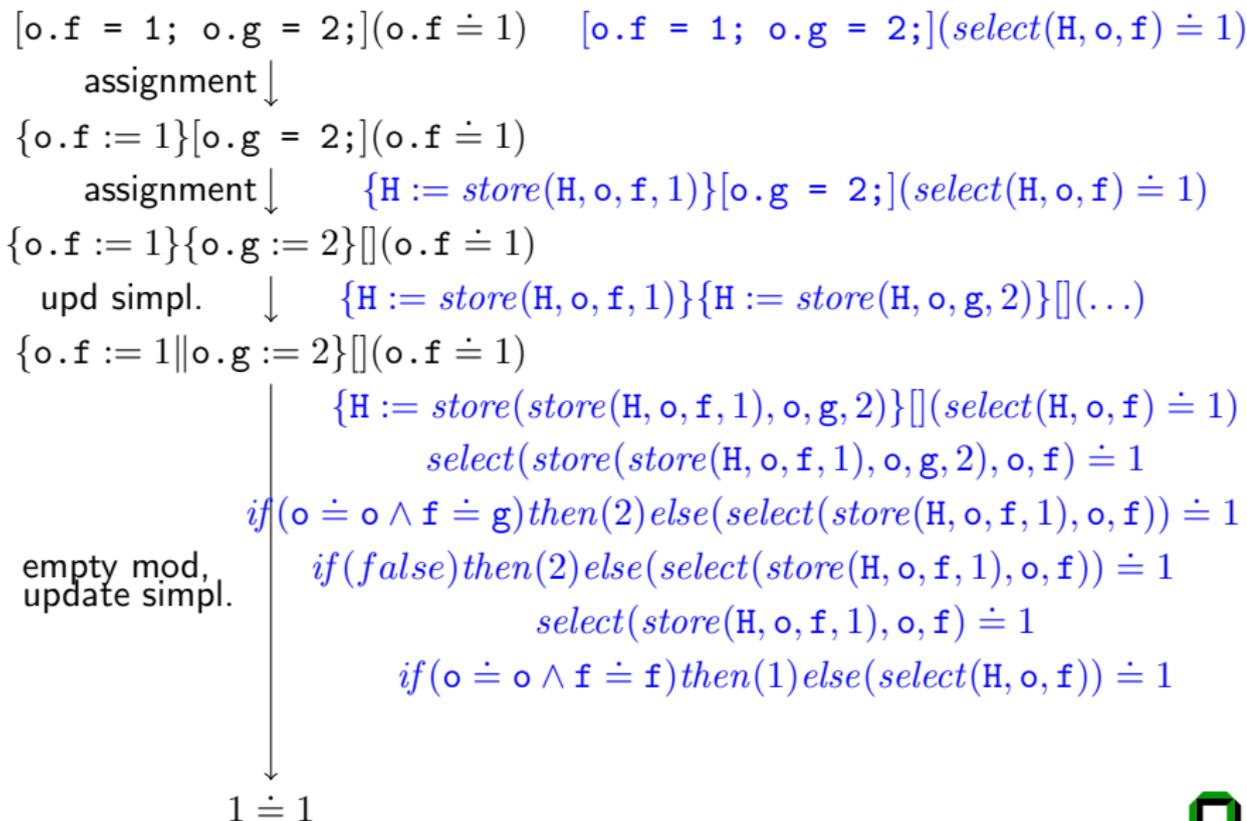
# A First Example (New Version)



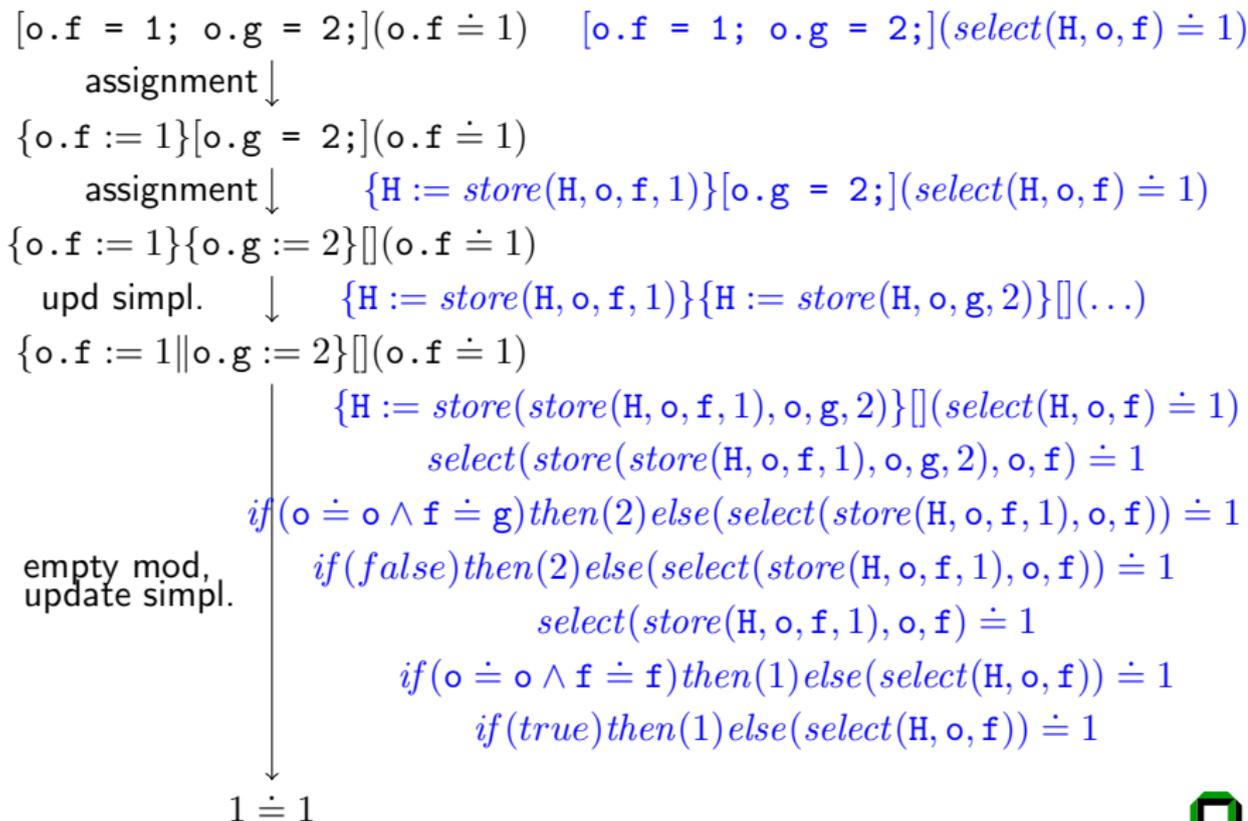
# A Frist Example (New Version)



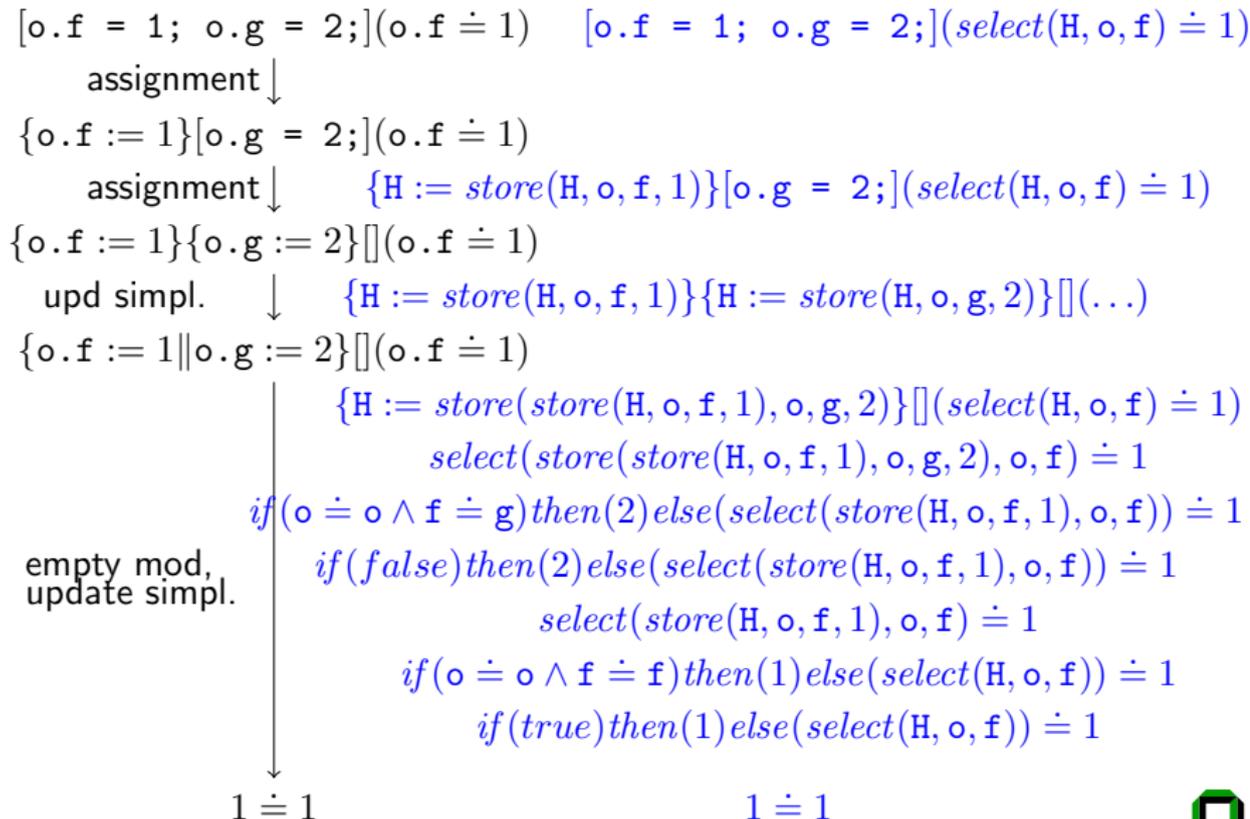
# A First Example (New Version)



# A Frist Example (New Version)



# A Frist Example (New Version)



# Benefits

- ▶ Modeling location dependent functions with the present heap representation grew into a monstrous task.



# Benefits

- ▶ Modeling location dependent functions with the present heap representation grew into a monstrous task.
- ▶ In particular modeling advanced concepts like *data groups* or *dynamic frames* poses serious problems.



# Benefits

- ▶ Modeling location dependent functions with the present heap representation grew into a monstrous task.
- ▶ In particular modeling advanced concepts like *data groups* or *dynamic frames* poses serious problems.
- ▶ Here is an example of a postcondition that formalizes purity of a method:

$$\forall Object\ o; \forall Field\ f; select(\mathbf{H}, o, f) \doteq select(H^{\text{@pre}}, o, f)$$

