

Towards imperative modules: reasoning about invariants and sharing of mutable state

David A. Naumann* and Mike Barnett

`naumann@cs.stevens-tech.edu`, `mbarnett@microsoft.com`

Stevens Institute of Technology and Microsoft Research

*Partially supported by NSF CCR-0208984 and Microsoft.

Programmer's intro: object invariants

```
class Subject {  
  private x,y: int := 0,1;  
  invariant  $\mathcal{I}(\text{self})$    where  $\mathcal{I}(o) = o.x < o.y$   
  method m() {  
    self.x := self.x+1;  
  
    self.y := self.y+1; } ... }
```

Programmer's intro: object invariants

```
class Subject {  
  private x,y: int := 0,1; obs: Observer :=...;  
  invariant  $\mathcal{I}(\text{self})$    where  $\mathcal{I}(o) = o.x < o.y$   
  method m() {  
    self.x := self.x+1;  
    obs.notify();  
    self.y := self.y+1; }   ... }
```

Programmer's intro: object invariants

```
class Subject {  
  private x,y: int := 0,1; obs: Observer :=...;  
  invariant  $\mathcal{I}(\text{self})$    where  $\mathcal{I}(o) = o.x < o.y$   
  method m() {  
    self.x := self.x+1;  
    obs.notify();  
    self.y := self.y+1; }    ... }  
class Observer {  
  z: Subject := ...;  
  method notify() { z.m(); }    ... }
```

Programmer's intro: object invariants

```
class Subject {  
  private x,y: int := 0,1; obs: Observer :=...;  
  invariant  $\mathcal{I}(\text{self})$    where  $\mathcal{I}(o) = o.x < o.y$   
  method m() {  
    self.x := self.x+1;  
    obs.notify();  
    self.y := self.y+1; } ... }  
class Observer {  
  z: Subject := ...;  
  method notify() { z.m(); } ... }
```

When should \mathcal{I} hold?

Programmer's intro (2): sharing

```
class Subject2 {  
  private x: Integer := new Integer(0);  
  private y: Integer := new Integer(1);  
  invariant  $\mathcal{I}(\text{self})$    where  $\mathcal{I}(o) = o.x.\text{val} < o.y.\text{val}$   
  method m() { self.x.incr(); self.y.incr(); }  
}
```

Programmer's intro (2): sharing

```
class Subject2 {  
  private x: Integer := new Integer(0);  
  private y: Integer := new Integer(1);  
  invariant  $\mathcal{I}(\text{self})$    where  $\mathcal{I}(o) = o.x.\text{val} < o.y.\text{val}$   
  method m() { self.x.incr(); self.y.incr(); }  
  method leak(): Integer { result := x; }    }  
class Main {  
  s: Subject2; i: Integer;  
  ... i := s.leak(); i.incr(); s.m() ... }
```

Programmer's intro (2): sharing

```
class Subject2 {  
  private x: Integer := new Integer(0);  
  private y: Integer := new Integer(1);  
  invariant  $\mathcal{I}(\text{self})$    where  $\mathcal{I}(o) = o.x.\text{val} < o.y.\text{val}$   
  method m() { self.x.incr(); self.y.incr(); }  
  method leak(): Integer { result := x; }    }  
class Main {  
  s: Subject2; i: Integer;  
  ... i := s.leak(); i.incr(); s.m() ... }
```

How to encapsulate not just fields but also referenced objects?

Logician's intro

$$\frac{\{P \wedge \mathcal{I}\} \text{ body } \{Q \wedge \mathcal{I}\} \quad \mathcal{I} \text{ is encapsulated for } m}{\{P\} \text{ call } m() \{Q\}}$$

Declaration: $m()\{\text{body}\}$. *Specification* \mathcal{P} , \mathcal{Q} involves public fields; \mathcal{I} depends on the internal representation.

Logician's intro

$$\frac{\{P \wedge \mathcal{I}\} \text{body} \{Q \wedge \mathcal{I}\} \quad \mathcal{I} \text{ is encapsulated for } m}{\{P\} \mathbf{call} \ m() \{Q\}}$$

Declaration: $m()\{\text{body}\}$. *Specification* \mathcal{P}, \mathcal{Q} involves public fields; \mathcal{I} depends on the internal representation.

$$\frac{\{P\} S \{Q\} \quad S \text{ does not interfere}^* \text{ with } \mathcal{I}}{\{P \wedge \mathcal{I}\} S \{Q \wedge \mathcal{I}\}}$$

Logician's intro

$$\frac{\{P \wedge \mathcal{I}\} \text{body} \{Q \wedge \mathcal{I}\} \quad \mathcal{I} \text{ is encapsulated for } m}{\{P\} \mathbf{call} \ m() \{Q\}}$$

Declaration: $m()\{\text{body}\}$. *Specification* \mathcal{P} , \mathcal{Q} involves public fields; \mathcal{I} depends on the internal representation.

$$\frac{\{P\} S \{Q\} \quad S \text{ does not interfere}^* \text{ with } \mathcal{I}}{\{P \wedge \mathcal{I}\} S \{Q \wedge \mathcal{I}\}}$$

* S does not write variables read in \mathcal{I} (hazard: aliasing)

* S does not update objects read in \mathcal{I} (hazard: heap sharing)

Logicians' intro (2)

$$\frac{\frac{\{R \quad \} x := E \{P \quad \}}{\{P \quad \} \mathbf{call} \ m \ {Q \quad \}} \quad \frac{\{P \wedge I\} \text{ body } \{Q \wedge I\}}{\{P \quad \} \mathbf{call} \ m \ {Q \quad \}}}{\{R \quad \} x := E ; \mathbf{call} \ m \ {Q \quad \}}$$

⋮

Logicians' intro (2)

$$\frac{\frac{\{R\} x := E \{P\} \quad \text{no interference}}{\{R \wedge I\} x := E \{P \wedge I\}} \quad \frac{\frac{\{P \wedge I\} \text{ body } \{Q \wedge I\}}{\{P \wedge I\} \mathbf{call} \ m \ \{Q \wedge I\}} \quad \vdots}{\{R \wedge I\} x := E ; \mathbf{call} \ m \ \{Q \wedge I\}} \quad \vdots$$

Logicians' intro (2)

$$\frac{\frac{\{R\} x := E \{P\} \quad \text{no interference}}{\{R \wedge I\} x := E \{P \wedge I\}} \quad \frac{\frac{\{P \wedge I\} \text{ body } \{Q \wedge I\}}{\{P \wedge I\} \mathbf{call} \ m \ \{Q \wedge I\}}}{\{R \wedge I\} x := E ; \mathbf{call} \ m \ \{Q \wedge I\}}}{\{R\} \text{ init}; \{R \wedge I\} x := E ; \mathbf{call} \ m \ \{Q \wedge I\} \Rightarrow \{Q\}}$$

Logicians' intro (2)

$$\begin{array}{c}
 \vdots \\
 \frac{\{R\} x := E \{P\} \quad \text{no interference}}{\{R \wedge I\} x := E \{P \wedge I\}} \qquad \frac{\{P \wedge I\} \text{ body } \{Q \wedge I\}}{\{P \wedge I\} \mathbf{call} \ m \ \{Q \wedge I\}} \\
 \hline
 \{R \wedge I\} x := E ; \mathbf{call} \ m \ \{Q \wedge I\} \\
 \vdots
 \end{array}$$

$$\{R\} \text{ init}; \{R \wedge I\} x := E ; \mathbf{call} \ m \ \{Q \wedge I\} \Rightarrow \{Q\}$$

How generalize to multiple instantiation, i.e., $\mathcal{I}(o)$ for all o ?

What about outcalls in `body`, i.e. method invocations on other objects?

How to express absence of interference, with heap sharing?

Contribution - a discipline

Problems:

- ◆ due to re-entrant callbacks, precondition $\mathcal{P} \wedge \mathcal{I}$ for body is unsound unless \mathcal{I} re-established before outcalls

Contribution - a discipline

Problems:

- ◆ due to re-entrant callbacks, precondition $\mathcal{P} \wedge \mathcal{I}$ for body is unsound unless \mathcal{I} re-established before outcalls
- ◆ need to protect $\mathcal{I}_{\text{Subject}}(o)$ from interference —by code in other classes and by other instances $p \neq o$ of Subject

Contribution - a discipline

Problems:

- ◆ due to re-entrant callbacks, precondition $\mathcal{P} \wedge \mathcal{I}$ for body is unsound unless \mathcal{I} re-established before outcalls
- ◆ need to protect $\mathcal{I}_{\text{Subject}}(o)$ from interference —by code in other classes and by other instances $p \neq o$ of Subject

Contribution: formalize and prove sound a solution using a

single everywhere-invariant, \mathcal{PI} . Rule:
$$\frac{\{\mathcal{PI} \wedge \mathcal{P}\} S \{Q\}}{\{\mathcal{P}\} S \{Q\}}.$$

Contribution - a discipline

Problems:

- ◆ due to re-entrant callbacks, precondition $\mathcal{P} \wedge \mathcal{I}$ for body is unsound unless \mathcal{I} re-established before outcalls
- ◆ need to protect $\mathcal{I}_{\text{Subject}}(o)$ from interference —by code in other classes and by other instances $p \neq o$ of Subject

Contribution: formalize and prove sound a solution using a

single everywhere-invariant, \mathcal{PI} . Rule:
$$\frac{\{\mathcal{PI} \wedge \mathcal{P}\} S \{Q\}}{\{\mathcal{P}\} S \{Q\}}.$$

Handles transfer and sharing of objects across encapsulation boundaries. Can use with standard logics.

Auxiliary field to make explicit when invariant holds:

*inv : **boolean** := false*

Maintain program invariant $PI : (\forall o \bullet o.inv \Rightarrow \mathcal{I}_{type(o)}(o))$

Auxiliary field to make explicit when invariant holds:

inv : **boolean** := false

Maintain program invariant $PI : (\forall o \bullet o.inv \Rightarrow \mathcal{I}_{type(o)}(o))$

class Subject { ...

invariant $\mathcal{I}_{Subject}(self)$ where $\mathcal{I}_{Subject}(o) = o.x < o.y$

method m() {

assert self.inv (* precondition *)

unpack self; (* self.inv := false *)

self.x := self.x+1; obs.notify(); self.y := self.y+1; (* $\mathcal{I}(self)$ *)

pack self; (* self.inv := true *) } ... }

class Main ... **method** notify() { **assert** z.inv ?; **z.m()**; }

Auxiliary field to make explicit when invariant holds:

inv : **boolean** := false

Maintain program invariant $PI : (\forall o \bullet o.inv \Rightarrow \mathcal{I}_{type(o)}(o))$

class Subject { ...

invariant $\mathcal{I}_{Subject}(self)$ where $\mathcal{I}_{Subject}(o) = o.x < o.y$

method m() {

assert self.inv (* precondition *)

unpack self; (* self.inv := false *)

self.x := self.x+1; obs.notify(); self.y := self.y+1; (* $\mathcal{I}(self)$ *)

pack self; (* self.inv := true *) } ... }

class Main ... **method** notify() { **assert** z.inv ?; z.m(); }

Absence of interf., as a precond.: $\{\neg v.inv \wedge PI\} v.f := E \{PI\}$

Auxiliary field to delimit heap dependence:

own : Object := null

Def: $o \preceq p$ iff either $o = p.\text{own}$ or $o \preceq p.\text{own}$.

Auxiliary field to delimit heap dependence:

own : Object := null

Def: $o \preceq p$ iff either $o = p.\text{own}$ or $o \preceq p.\text{own}$.

Def: \mathcal{I}_C is admissible iff

when $\mathcal{I}_C(o)$ depends on $p.f$ then either $o = p$ or $o \preceq p$.

Auxiliary field to delimit heap dependence:

own : Object := null

Def: $o \preceq p$ iff either $o = p.\text{own}$ or $o \preceq p.\text{own}$.

Def: \mathcal{I}_C is admissible iff

when $\mathcal{I}_C(o)$ depends on $p.f$ then either $o = p$ or $o \preceq p$.

Absence of interference, as a precondition:

$\{\neg v.\text{inv} \wedge (\forall o \bullet o \preceq v \Rightarrow \neg o.\text{inv}) \wedge \mathcal{PI}\} v.f := E \{\mathcal{PI}\}$

Ownership provides stateful encapsulation: $\neg o.\text{inv}$ means control is inside the boundary for o .

Last auxiliary field for ownership discipline:

`com : boolean := false`

$\mathcal{PI} : (\forall o \bullet o.inv \Rightarrow \mathcal{I}_{type(o)}(o)) \wedge$

$(\forall o \bullet o.inv \Rightarrow (\forall p \bullet p.own = o \Rightarrow p.com)) \wedge$

$(\forall o \bullet o.com \Rightarrow o.inv)$

Last auxiliary field for ownership discipline:

com : **boolean** := false

$\mathcal{PI} : (\forall o \bullet o.inv \Rightarrow \mathcal{I}_{type(o)}(o)) \wedge$

$(\forall o \bullet o.inv \Rightarrow (\forall p \bullet p.own = o \Rightarrow p.com)) \wedge$

$(\forall o \bullet o.com \Rightarrow o.inv)$

Absence of interf., as a precondition: $\{\neg v.inv \wedge \mathcal{PI}\} v.f := E \{\mathcal{PI}\}$

Last auxiliary field for ownership discipline:

com : **boolean** := false

$\mathcal{PI} : (\forall o \bullet o.inv \Rightarrow \mathcal{I}_{type(o)}(o)) \wedge$

$(\forall o \bullet o.inv \Rightarrow (\forall p \bullet p.own = o \Rightarrow p.com)) \wedge$

$(\forall o \bullet o.com \Rightarrow o.inv)$

Absence of interf., as a precond.: $\{ \neg v.inv \wedge \mathcal{PI} \} v.f := E \{ \mathcal{PI} \}$

*Precondition and effect of **unpack** E:*

assert E.inv \wedge \neg E.com;

E.inv := false; **forall** o **with** o.own = E **do** o.com := false;

Last auxiliary field for ownership discipline:

com : **boolean** := false

$\mathcal{PI} : (\forall o \bullet o.inv \Rightarrow \mathcal{I}_{type(o)}(o)) \wedge$

$(\forall o \bullet o.inv \Rightarrow (\forall p \bullet p.own = o \Rightarrow p.com)) \wedge$

$(\forall o \bullet o.com \Rightarrow o.inv)$

Absence of interf., as a precondition: $\{ \neg v.inv \wedge \mathcal{PI} \} v.f := E \{ \mathcal{PI} \}$

*Precondition and effect of **unpack** E:*

assert E.inv \wedge \neg E.com;

E.inv := false; **forall** o **with** o.own = E **do** o.com := false;

*Precondition and effect of **pack** E:*

assert \neg E.inv \wedge $\mathcal{I}_{type(E)}(E)$;

E.inv := true; **forall** o **with** o.own = E **do** o.com := true;

Ownership transfer

$\{\neg v.\text{inv} \wedge \neg E.\text{inv} \wedge \mathcal{PI}\} v.\text{own} := E \{\mathcal{PI}\}$

*Special command **setowner** to highlight that it only manipulates auxiliary state (like **unpack/pack**.*

State-based encapsulation (vs. type systems):

- ◆ *avoids restriction on existence or reading of references*
- ◆ *allows transfer of objects across boundaries*

Main result: stateful encapsulation I.

Def: S is *properly annotated* iff each **pack**, **unpack**, **setowner**, and field update has stipulated precondition.

Theorem: $\{PI\} S \{PI\}$ for any properly annotated S

Rule:
$$\frac{\{PI \wedge P\} S \{Q\}}{\{P\} S \{Q\}}$$

Proof: using a straightforward denotational semantics for a sequential language with mutually recursive class declarations and methods etc.

Stateful encapsulation II.

A List owns its nodes. A node does not own its neighbors.

```
class List {  
  head: ListNode;  
  invariant self.head=null  $\vee$  self.head.prev=null; ... }  
class ListNode {  
  next, prev: ListNode;  
  invariant self.next=null  
     $\vee$  (self.next.prev=self  $\wedge$  self.next.own=self.own); ... }
```


Stateful encapsulation II.

A List owns its nodes. A node does not own its neighbors.

```
class List {  
  head: ListNode;  
  invariant self.head=null  $\vee$  self.head.prev=null;  ... }  
class ListNode {  
  next, prev: ListNode;  
  invariant self.next=null  
     $\vee$  (self.next.prev=self  $\wedge$  self.next.own=self.own);  ... }
```

Decentralized invariants express acyclicity without induction. Well behaved interaction but not ownership.

Absence of interference, as a precondition:

$\{\neg v.\text{inv} \wedge$

$(\forall o \bullet \text{“}\mathcal{I}_C(o) \text{ depends on } v.f\text{”} \Rightarrow \neg o.\text{inv} \vee \mathcal{I}_C(o)[E/v.f])\}$

$v.f := E$

Absence of interference, as a precondition:

$\{\neg v.\text{inv} \wedge$

$(\forall o \bullet \text{“}\mathcal{I}_C(o) \text{ depends on } v.f\text{”} \Rightarrow \neg o.\text{inv} \vee \mathcal{I}_C(o)[E/v.f])\}$

$v.f := E$

Auxiliary field for friendship discipline:

$\text{deps} : \text{set of Object} := \emptyset$

$(\forall o \bullet o \in v.\text{deps} \Rightarrow \neg o.\text{inv} \vee \mathcal{I}_C(o)[E/v.f])$

Absence of interference, as a precondition:

$\{\neg v.\text{inv} \wedge$

$(\forall o \bullet \text{“}\mathcal{I}_C(o) \text{ depends on } v.f\text{”} \Rightarrow \neg o.\text{inv} \vee \mathcal{I}_C(o)[E/v.f])\}$

$v.f := E$

Auxiliary field for friendship discipline:

deps : set of Object := \emptyset

$(\forall o \bullet o \in v.\text{deps} \Rightarrow \neg o.\text{inv} \vee \mathcal{I}_C(o)[E/v.f])$

*Admissibility: when $\mathcal{I}_C(o)$ depends on $p.f$ then either $o = p$, $o \preceq p$, or **$p = o.g$** for some declared pivot **g** and **$o \in p.\text{deps}$***

Absence of interference, as a precondition:

$\{\neg v.\text{inv} \wedge$

$(\forall o \bullet \text{“}\mathcal{I}_C(o) \text{ depends on } v.f\text{”} \Rightarrow \neg o.\text{inv} \vee \mathcal{I}_C(o)[E/v.f])\}$

$v.f := E$

Auxiliary field for friendship discipline:

deps : set of Object := \emptyset

$(\forall o \bullet o \in v.\text{deps} \Rightarrow \neg o.\text{inv} \vee \mathcal{I}_C(o)[E/v.f])$

*Admissibility: when $\mathcal{I}_C(o)$ depends on $p.f$ then either $o = p$, $o \preceq p$, or **$p = o.g$** for some declared pivot **g** and **$o \in p.\text{deps}$***

Abstract from $\mathcal{I}_C(o)[E/v.f]$ as $\mathcal{U}_C(o, v, E)$.

Obligation: $\{\mathcal{I}_C(\text{self}) \wedge \mathcal{U}(\text{self}, g, \text{val})\}$ self.g.f := val $\{\mathcal{I}_C(\text{self})\}$

Conclusion

- ◆ Discipline for control of dependence for object invariants. Controls use of pointers rather than their existence.
- ◆ Handles difficult design patterns that are common in practice. No restrictions on heap structure.
- ◆ No commitment to particular program logic or verification system.
- ◆ Uses verification conditions; not special type annotation but not fully automated.

Future work

- ◆ comparison with Separation Logic:

$$\frac{\{P\} m \{Q\} \vdash \{P'\} S \{Q'\}}{\{P * I\} \text{ body } \{Q * I\} \vdash \{P' * I\} S \{Q' * I\}}$$

- ◆ multi-class patterns
- ◆ implementation and case studies —Boogie project
- ◆ implementation and full logic —Pierik and de Boer
- ◆ representation independence —Banerjee and Naumann
- ◆ friends and subclassing —Barnett and Naumann
- ◆ machine check soundness proof —Naumann

References

- ◆ Barnett, DeLine, Fähndrich, Leino, Wolfram Schulte: *Verification of object-oriented programs with invariants* (Journal of Object Technology '04)
- ◆ Leino and Müller: *Object invariants in dynamic contexts* (ECOOP'04)
- ◆ Barnett and Naumann: *Friends need a bit more: maintaining invariants over shared state* (MPC'04)
- ◆ Pierik, Clarke, de Boer: *Creational invariants* (FTfJP'04)
- ◆ O'Hearn, Yang, Reynolds: *Separation and Information Hiding* (POPL'04)