

The Rule of Constancy (Derived Frame Rule)

- The following derived rule is used on the next slide

The rule of constancy

$$\frac{\vdash \{P\} C \{Q\}}{\vdash \{P \wedge R\} C \{Q \wedge R\}}$$

where no variable assigned to in C occurs in R

- Outline of derivation
 - prove $\{R\} C \{R\}$ by induction on C
 - then use Specification Conjunction
- Assume C doesn't modify V and $\vdash \{P\} C \{P[V+1/V]\}$ then:
 - $\vdash \{P \wedge V=v\} C \{P[V+1/V] \wedge V=v\}$ (assumption + constancy rule)
 - $\vdash \{P[V+1/V] \wedge V=v\} V := V+1 \{P \wedge V=v+1\}$ (assign. ax + pre. streng.)
 - $\vdash \{P \wedge V=v\} C; V := V+1 \{P \wedge V=v+1\}$ (sequencing)
- So $C; V := V+1$ has P as an invariant and increments V

Towards the FOR-Rule ✓

- If $e_1 \leq e_2$ the FOR-command is equivalent to:

BEGIN VAR V ; $V := e_1$; ... C ; $V := V+1$; ... $V := e_2$; C END

- Assume **C doesn't modify V** and **$\vdash \{P\} C \{P[V+1/V]\}$**

- Hence:

$\vdash \{P[e_1/V]\} V := e_1 \{P \wedge V=e_1\}$ (assign. ax + pre. streng.)

⋮

$\vdash \{P \wedge V=v\} C ; V := V+1 \{P \wedge V=v+1\}$ (last slide; $V = e_1, e_1+1, \dots, e_2-1$)

⋮

$\vdash \{P \wedge V=v\} C ; V := V+1 \{P \wedge V=e_2+1\}$

$\vdash \{P \wedge V=e_2\} C \{P[V+1/V] \wedge V=e_2\}$ (assign. ax + assumption + constancy)

$\vdash \{P \wedge V=e_2\} C \{P[e_2+1/V]\}$ (post. weak.)

- Hence by the sequencing and block rules

$$\frac{\vdash \{P\} C \{P[V+1/V]\}}{\vdash \{P[e_1/V]\} \text{BEGIN VAR } V ; V := e_1 ; \dots C ; V := V+1 ; \dots V := e_2 ; C \text{ END} \{P[e_2+1/V]\}}$$

The FOR-Rule

- To rule out the problems that arise when the controlled variable or variables in the bounds expressions, are changed by the body, we simply impose a side condition on the rule that stipulates that it cannot be used in these situations

The FOR-rule

$$\frac{\vdash \{P \wedge (E_1 \leq V) \wedge (V \leq E_2)\} C \{P[V+1/V]\}}{\vdash \{P[E_1/V] \wedge (E_1 \leq E_2)\} \text{ FOR } V := E_1 \text{ UNTIL } E_2 \text{ DO } C \{P[E_2+1/V]\}}$$

where neither V , nor any variable occurring in E_1 or E_2 , is assigned to in the command C .

- Note $(E_1 \leq V) \wedge (V \leq E_2)$ in precondition of rule hypothesis
 - added to strengthen rule to allow proofs to use facts about V 's range of values
- Can be tricky to think up P

Comment on the FOR-Rule

- The FOR-rule does not enable anything to be deduced about FOR-commands whose body assigns to variables in the bounds expressions
- This precludes such assignments being used if commands are to be reasoned about
- Only defining rules of inference for non-tricky uses of constructs motivates writing programs in a perspicuous manner
- It is possible to devise a rule that does cope with assignments to variables in bounds expressions
- Consider the rule below (e_1, e_2 are fresh auxiliary variables):

$$\frac{\vdash \{P \wedge (e_1 \leq V) \wedge (V \leq e_2)\} C \{P[V+1/V]\}}{\vdash \{P[E_1/V] \wedge (E_1 \leq E_2) \wedge (E_1 = e_1) \wedge (E_2 = e_2)\} \text{ FOR } V := E_1 \text{ UNTIL } E_2 \text{ DO } C \{P[e_2+1/V]\}}$$

The FOR-axiom

- To cover the case when $E_2 < E_1$, we need the FOR-axiom below

The FOR-axiom

$$\vdash \{P \wedge (E_2 < E_1)\} \text{ FOR } V := E_1 \text{ UNTIL } E_2 \text{ DO } C \{P\}$$

- This says that when E_2 is less than E_1 the FOR-command has no effect

Ensuring Soundness

- It is clear from the discussion of the FOR-rule that it is not always straightforward to devise correct rules of inference
- It is important that the axioms and rules be sound. There are two approaches to ensure this
 - (i) define the language by the axioms and rules of the logic
 - (ii) prove that the logic is sound for the language
- Approach (i) is called *axiomatic semantics*
 - the idea is to *define* the semantics of the language by requiring that it make the axioms and rules of inference true
 - it is then up to implementers to ensure that the logic matches the language
- Approach (ii) is proving soundness of the logic

Axiomatic Semantics

- One snag with axiomatic semantics is that most existing languages have already been defined in some other way
 - usually by informal and ambiguous natural language statements
- The other snag with axiomatic semantics is that by Clarke's Theorem it is known to be impossible to devise relatively complete Floyd-Hoare logics for languages with certain constructs
 - it could be argued that this is not a snag at all but an advantage, because it forces programming languages to be made logically tractable
- An example of a language defined axiomatically is Euclid

7.1. (module rule)

$$(1) \quad Q \supset Q0(A/t),$$

$$(2) \quad P1\{\mathbf{const} \ K; \mathbf{var} \ V; S_4\} Q4(A/t) \wedge Q,$$

$$(3) \quad P2(A/t) \wedge Q\{S_2\} Q2(A/t) \wedge Q,$$

$$(4) \quad \exists g1(P3(A/t) \wedge Q\{S_3\} Q3(A/t) \wedge g = g1(A, c, d)),$$

$$(5) \quad \exists g(P3(A/t) \wedge Q \supset Q3(A/t)),$$

$$(6) \quad P6(A/t) \wedge Q\{S_6\} Q1,$$

$$(7) \quad P \supset P1(a/c),$$

$$(8.1) \quad [Q0(a/c, x/t, x'/t') \supset (P2(x/t, x'/t', a2/x2, e2/c2, a/c) \wedge \\ (Q2(x2\#/t, x'/t', a2\#/x2, e2/c2, a/c, y2\#/y2, a2/x2', y2/y2') \supset \\ R1(x2\#/x, a2\#/a2, y2\#/y2))) \{x. p(a2, e2)\} R1 \wedge Q0(a/c, x/t, x'/t'),$$

$$(8.2) \quad (Q0(a/c, x/t) \supset P3(x/t, a3/c3, a/c)) \supset \\ Q3(x/t, a3/c3, a/c, f(a3, d3)/g) \wedge Q0(a/c, x/t),$$

$$(8.3) \quad P1(a/c) \wedge (Q4(x4\#/t, x'/t', a/c, y4\#/y4, y4/y4') \supset R4(x4\#/x, y4\#/y4)) \\ \{x. \text{Initially}\} R4 \wedge Q0(a/c, x/t, x'/t'),$$

$$(8.4) \quad (Q0(a/c, x/t, x'/t') \supset P6(x/t, x'/t', a/c)) \wedge (Q1(a/c, y6\#/y6, y6/y6') \supset \\ R(y6\#/y6)) \{x. \text{Finally}\} R]$$

⊢

$$(8.5) \quad P(x\#/x) \{x. \text{Initially}; S; x. \text{Finally}\} R(x\#/x)$$

$$P\{\mathbf{var} \ x: T(a); S\} R \wedge Q1$$

Array assignments

- **Syntax:** $V(E_1) := E_2$
- **Semantics:** the state is changed by assigning the value of the term E_2 to the E_1 -th component of the array variable V
- **Example:** $A(X+1) := A(X)+2$
 - if the the value of X is x
 - and the value of the x -th component of A is n
 - then the value stored in the $(x+1)$ -th component of A becomes $n+2$

Naive Array Assignment Axiom Fails

- The axiom

$$\vdash \{P[E_2/A(E_1)]\} A(E_1) := E_2 \{P\}$$

doesn't work

- Take $P \equiv 'X=Y \wedge A(Y)=0'$, $E_1 \equiv 'X'$, $E_2 \equiv '1'$
 - since $A(X)$ does not occur in P
 - it follows that $P[1/A(X)] = P$
 - hence the axiom yields: $\vdash \{X=Y \wedge A(Y)=0\} A(X) := 1 \{X=Y \wedge A(Y)=0\}$
- Must take into account possibility that changes to $A(X)$ may change $A(Y)$, $A(Z)$ etc
 - since X might equal Y , Z etc (i.e. **aliasing**)
- Related to the *Frame Problem in AI*

Reasoning About Arrays

- The naive array assignment axiom

$$\vdash \{P[E_2/A(E_1)]\} A(E_1) := E_2 \{P\}$$

does not work: changes to $A(X)$ may also change $A(Y)$, $A(Z)$, ...

- The solution, due to Hoare, is to treat an array assignment

$$A(E_1) := E_2$$

as an ordinary assignment

$$A := A\{E_1 \leftarrow E_2\}$$

where the term $A\{E_1 \leftarrow E_2\}$ denotes an array identical to A , except that the E_1 -th component is changed to have the value E_2

Array Assignment axiom

- Array assignment is a special case of ordinary assignment

$$A := A\{E_1 \leftarrow E_2\}$$

- Array assignment axiom just ordinary assignment axiom

$$\vdash \{P[A\{E_1 \leftarrow E_2\}/A]\} A := A\{E_1 \leftarrow E_2\} \{P\}$$

- Thus:

The array assignment axiom

$$\vdash \{P[A\{E_1 \leftarrow E_2\}/A]\} A(E_1) := E_2 \{P\}$$

Where A is an array variable, E_1 is an integer valued expression, P is any statement and the notation $A\{E_1 \leftarrow E_2\}$ denotes the array identical to A , except that $A(E_1) = E_2$.

Array Axioms

- In order to reason about arrays, the following axioms, which define the meaning of the notation $A\{E_1 \leftarrow E_2\}$, are needed

The array axioms

$$\vdash A\{E_1 \leftarrow E_2\}(E_1) = E_2$$

$$\vdash E_1 \neq E_3 \Rightarrow A\{E_1 \leftarrow E_2\}(E_3) = A(E_3)$$

- Second of these is a *Frame Axiom*
 - don't confuse with Frame Rule of Separation Logic (later)
 - “frame” is a rather overloaded word!

New Topic: Separation logic

- One of several competing methods for reasoning about pointers
- Details took 30 years to evolve
- Shape predicates due to Rod Burstall in the 1970s
- Separation logic: by O'Hearn, Reynolds and Yang around 2000
- Several partially successful attempts before separation logic
- Very active research area
 - QMUL, UCL, Cambridge, Harvard, Princeton, Yale
 - Microsoft
 - startup Monoidics bought by Facebook

Pointers and the state

- So far the state just determined the values of variables
 - values assumed to be numbers
 - preconditions and postconditions are first-order logic statements
 - state same as a valuation $s : Var \rightarrow Val$
- To model pointers – e.g. as in C – add *heap* to state
 - heap maps *locations* (pointers) to their contents
 - *store* maps variables to values (previously called state)
 - contents of locations can be locations or values

$$\begin{array}{ccccccc} X & \mapsto & l_1 & \mapsto & l_2 & \mapsto & v \\ & \text{store} & & \text{heap} & & \text{heap} & \end{array}$$

Heap semantics

$$Store = Var \rightarrow Val$$

(assume $Num \subseteq Val$, $nil \in Val$ and $nil \notin Num$)

$$Heap = Num \rightarrow_{fin} Val$$

$$State = Store \times Heap$$

- **Note:** store also called *stack* or *environment*; heap also called *store*

Adding pointer operations to our language

Expressions:

$E ::= N \mid V \mid E_1 + E_2 \mid E_1 - E_2 \mid E_1 \times E_2 \mid \dots$

Boolean expressions:

$B ::= T \mid F \mid E_1 = E_2 \mid E_1 \leq E_2 \mid \dots$

commands:

$C ::= V := E$	value assignments
$V := [E]$	fetch assignments
$[E_1] := E_2$	heap assignments (heap mutation)
$V := \text{cons}(E_1, \dots, E_n)$	allocation assignments
$\text{dispose}(E)$	pointer disposal
$C_1 ; C_2$	sequences
$\text{IF } B \text{ THEN } C_1 \text{ ELSE } C_2$	conditionals
$\text{WHILE } B \text{ DO } C$	while commands

Pointer manipulation constructs and faulting

- Commands executed in a state (s, h)
- Reading, writing or disposing pointers might *fault*
- Fetch assignments: $V := [E]$
 - evaluate E to get a location l
 - fault if l is not in the heap
 - otherwise assign contents of l in heap to the variable V
- Heap assignments: $[E_1] := E_2$
 - evaluate E_1 to get a location l
 - fault if the l is not in the heap
 - otherwise store the value of E_2 as the new contents of l in the heap
- Pointer disposal: $\text{dispose}(E)$
 - evaluate E to get a pointer l (a number)
 - fault if l is not in the heap
 - otherwise remove l from the heap

Allocation assignments

- Allocation assignments: $V := \text{cons}(E_1, \dots, E_n)$
 - choose n consecutive locations that are not in the heap, say $l, l+1, \dots$
 - extend the heap by adding $l, l+1, \dots$ to it
 - assign l to the variable V in the store
 - make the values of E_1, E_2, \dots be the new contents of $l, l+1, \dots$ in the heap
- Allocation assignments never fault
- Allocation assignments are *non-deterministic*
 - any suitable $l, l+1, \dots$ not in the heap can be chosen
 - always exists because the heap is finite

Example (different from the background reading) ✓

$X := \text{cons}(0, 1, 2); [X] := Y+1; [X+1] := Z; Y := [Y+Z]$

- $X := \text{cons}(0, 1, 2)$ allocates three new pointers, say $l, l+1, l+2$
 - l initialised with contents 0, $l+1$ with 1 and $l+2$ with 2
 - variable X is assigned l as its value in store
- $[X] := Y+1$ changes the contents of l
 - l gets value of $Y+1$ as new contents in heap
- $[X+1] := Z$ changes the contents of $l+1$
 - $l+1$ gets the value of Z as new contents in heap
- $Y := [Y+Z]$ changes the value of Y in the store
 - Y assigned in the store the contents of $Y+Z$ in the heap
 - faults if the value of $Y+Z$ is not in the heap

Local Reasoning and Separation Logic

- Want to just reason about just those locations being modified
 - assume all other locations unchanged
- Solution: separation logic
 - **small** and **forward** assignment axioms + **separating conjunction**
 - **small** means just applies to fragment of heap (*footprint*)
 - **forward** means Floyd-style forward rules that support symbolic execution
 - **non-faulting semantics** of Hoare triples
 - symbolic execution used by tools like smallfoot
 - **separating conjunction** solves frame problem - like rule of constancy for heap
- Need new kinds of assertions to state separation logic axioms

Sneak preview of the Frame Rule

The frame rule

$$\frac{\vdash \{P\} C \{Q\}}{\vdash \{P \star R\} C \{Q \star R\}}$$

where no variable modified by C occurs free in R .

- Separating conjunction $P \star Q$
 - heap can be split into two disjoint components
 - P is true of one component and Q of the other
 - \star is commutative and associative

Local Reasoning and Separation Logic

- Want to just reason about just those locations being modified
 - assume all other locations unchanged
- Solution: separation logic
 - **small** and **forward** assignment axioms + **separating conjunction**
 - **small** means just applies to fragment of heap (*footprint*)
 - **forward** means Floyd-style forward rules that support symbolic execution
 - **non-faulting semantics** of Hoare triples
 - symbolic execution used by tools like smallfoot
 - **separating conjunction** solves frame problem - like rule of constancy for heap
- Need new kinds of assertions to state separation logic axioms

Separation logic assertions: emp

- emp is an atomic statement of separation logic
- emp is true iff the heap is empty
- The semantics of emp is:
$$\text{emp}(s, h) \Leftrightarrow h = \{\}$$
(where $\{\}$ is the empty heap)
- Abbreviation: $E_1 \dot{=} E_2 =_{def} (E_1 = E_2) \wedge \text{emp}$
- From the semantics: $(E_1 \dot{=} E_2)(s, h) \Leftrightarrow E_1(s) = E_2(s) \wedge h = \{\}$
- $E_1 = E_2$ is independent of the heap and only depends on the store
- Semantics of $E_1 = E_2$ is:
$$(E_1 = E_2)(s, h) \Leftrightarrow E_1(s) = E_2(s)$$

no constraint on the heap – any h will do

Separation logic: small axioms and faulting

- One might expect a heap assignment axiom to entail:

$$\vdash \{\text{T}\} [0] := 0 \{0 \mapsto 0\}$$

i.e. after executing $[0] := 0$ the contents of location 0 in the heap is 0

- Recall the sneak preview of the frame rule:

The frame rule

$$\frac{\vdash \{P\} C \{Q\}}{\vdash \{P \star R\} C \{Q \star R\}}$$

where no variable modified by C occurs free in R .

- Taking R to be the points-to statement $0 \mapsto 1$ yields:

$$\vdash \{\text{T} \star 0 \mapsto 1\} [0] := 0 \{0 \mapsto 0 \star 0 \mapsto 1\}$$

something is wrong with the conclusion!

- Solution: define Hoare triple so $\vdash \{\text{T}\} [0] := 0 \{0 \mapsto 0\}$ is not true

Non-faulting interpretation of Hoare triples

- The *non-faulting semantics* of Hoare triples $\{P\} C \{Q\}$ is:

if P holds then

- (i) executing C doesn't fault **and**
- (ii) if C terminates then Q holds

$$\models \{P\} C \{Q\} = \forall s h. P(s, h) \Rightarrow \neg(C(s, h)\text{fault}) \wedge \forall s' h'. C(s, h)(s', h') \Rightarrow Q(s', h')$$

- Now $\vdash \{T\} [0] := 0 \{0 \mapsto 0\}$ is not true as $([0] := 0)(s, \{\})\text{fault}$
- Recall the sneak preview of the frame rule:

The frame rule

$$\frac{\vdash \{P\} C \{Q\}}{\vdash \{P \star R\} C \{Q \star R\}}$$

where no variable modified by C occurs free in R .

- So can't use frame rule to get $\vdash \{T \star 0 \mapsto 1\} [0] := 0 \{0 \mapsto 0 \star 0 \mapsto 1\}$

Store assignment axiom

Store assignment axiom

$$\vdash \{V \doteq v\} V := E \{V \doteq E[v/V]\}$$

where v is an auxiliary variable not occurring in E .

- $E_1 \doteq E_2$ means value of E_1 and E_2 equal in the store **and** heap is empty
- In Hoare logic (no heap) this is equivalent to the assignment axiom

$$\vdash \{V=v\} V := E \{V=E[v/V]\}$$

store assign. ax.

$$\vdash \{V=v \wedge Q[E[v/V]/V]\} V := E \{V=E[v/V] \wedge Q[E[v/V]/V]\}$$

rule of constancy

$$\vdash \{\exists v. V=v \wedge Q[E[v/V]/V]\} V := E \{\exists v. V=E[v/V] \wedge Q[E[v/V]/V]\}$$

exists introduction

$$\vdash \{\exists v. V=v \wedge Q[E[V/V]/V]\} V := E \{\exists v. V=E[v/V] \wedge Q[V/V]\}$$

predicate logic

$$\vdash \{\exists v. V=v \wedge Q[E/V]\} V := E \{\exists v. V=E[v/V] \wedge Q\}$$

$[V/V]$ is identity

$$\vdash \{(\exists v. V=v) \wedge Q[E/V]\} V := E \{(\exists v. V=E[v/V]) \wedge Q\}$$

predicate logic: v not in E

$$\vdash \{\top \wedge Q[E/V]\} V := E \{(\exists v. V=E[v/V]) \wedge Q\}$$

predicate logic

$$\vdash \{Q[E/V]\} V := E \{Q\}$$

rules of consequence

- Separation logic: exists introduction valid, rule of constancy invalid

Fetch assignment axiom

Fetch assignment axiom

$$\vdash \{(V = v_1) \wedge E \mapsto v_2\} V := [E] \{(V = v_2) \wedge E[v_1/V] \mapsto v_2\}$$

where v_1, v_2 are auxiliary variables not occurring in E .

- Precondition guarantees the assignment doesn't fault
- V is assigned the contents of E in the heap
- Small axiom: precondition and postcondition specify singleton heap
- If neither V nor v occur in E then the following holds:

$$\vdash \{E \mapsto v\} V := [E] \{(V = v) \wedge E \mapsto v\}$$

(proof: instantiate v_1 to V and v_2 to v and then simplify)

Heap assignment axiom

Heap assignment axiom (heap mutation)

$$\vdash \{E \mapsto _ \} [E] := F \{E \mapsto F \}$$

- Precondition guarantees the assignment doesn't fault
- Contents of E in heap is updated to be value of F
- Small axiom: precondition and postcondition specify singleton heap

Pointer allocation

Allocation assignment axiom

$$\vdash \{V \doteq v\} V := \text{cons}(E_1, \dots, E_n) \{V \mapsto E_1[v/V], \dots, E_n[v/V]\}$$

where v is an auxiliary variable not equal to V or occurring in E_1, \dots, E_n

- Never faults

- If V doesn't occur in E_1, \dots, E_n then:

$$\vdash \{V \doteq v\} V := \text{cons}(E_1, \dots, E_n) \{V \mapsto E_1[v/V], \dots, E_n[v/V]\} \quad \text{alloc. assign. ax}$$

$$\vdash \{V \doteq v\} V := \text{cons}(E_1, \dots, E_n) \{V \mapsto E_1, \dots, E_n\} \quad V \text{ not in } E_i \text{ assump.}$$

$$\vdash \{\exists v. V \doteq v\} V := \text{cons}(E_1, \dots, E_n) \{\exists v. V \mapsto E_1, \dots, E_n\} \quad \text{exists intro.}$$

$$\vdash \{\exists v. V = v \wedge \text{emp}\} V := \text{cons}(E_1, \dots, E_n) \{\exists v. V \mapsto E_1, \dots, E_n\} \quad \text{definition of } \doteq$$

$$\vdash \{\text{emp}\} V := \text{cons}(E_1, \dots, E_n) \{V \mapsto E_1, \dots, E_n\} \quad \text{predicate logic}$$

- Which is a derivation of:

Derived allocation assignment axiom

$$\vdash \{\text{emp}\} V := \text{cons}(E_1, \dots, E_n) \{V \mapsto E_1, \dots, E_n\}$$

where V doesn't occur in E_1, \dots, E_n .

Pointer deallocation

Dispose axiom

$$\vdash \{E \mapsto _ \} \text{dispose}(E) \{ \text{emp} \}$$

- Attempting to deallocate a pointer not in the heap faults
- Small axiom: singleton precondition heap, empty postcondition heap
- Sanity checking example proof:

$$\vdash \{E_1 \mapsto _ \} \text{dispose}(E_1) \{ \text{emp} \}$$

dispose axiom

$$\vdash \{ \text{emp} \} V := \text{cons}(E_2) \{ V \mapsto E_2 \}$$

derived allocation assignment axiom

$$\vdash \{E_1 \mapsto _ \} \text{dispose}(E_1); V := \text{cons}(E_2) \{ V \mapsto E_2 \}$$

sequencing rule

Compound command rules

- Following rules apply to both Hoare logic and separation logic

The sequencing rule

$$\frac{\vdash \{P\} C_1 \{Q\}, \quad \vdash \{Q\} C_2 \{R\}}{\vdash \{P\} C_1; C_2 \{R\}}$$

The conditional rule

$$\frac{\vdash \{P \wedge S\} C_1 \{Q\}, \quad \vdash \{P \wedge \neg S\} C_2 \{Q\}}{\vdash \{P\} \text{IF } S \text{ THEN } C_1 \text{ ELSE } C_2 \{Q\}}$$

The WHILE-rule

$$\frac{\vdash \{P \wedge S\} C \{P\}}{\vdash \{P\} \text{WHILE } S \text{ DO } C \{P \wedge \neg S\}}$$

- For separation logic, need to think about faulting