

Mechanising Staged Logic

Darius Foo, Yahui Song, Wei-Ngan Chin

SG PL Summit

4 December 2024

Staged logic

- A new, alternative program logic formulation
- Automated (SMT-based) verification
- Effectful higher-order programs

Staged Specification Logic for Verifying Higher-Order Imperative Programs

Darius Foo(✉)^[0000-0002-3279-5827], Yahui Song^[0000-0002-9760-5895], and
Wei-Ngan Chin^[0000-0002-9660-5682]

School of Computing, National University of Singapore, Singapore
{dariusf,yahuis,chinwn}@comp.nus.edu.sg

An effectful higher-order program

```
let x = ref [] in  
foldr (fun c t -> x := c :: !x; c + t) xs 0
```

A specification we would like to give it

$$\forall x\ a, \{x \mapsto a\}$$
$$\text{foldr } (\text{fun } c\ t \rightarrow x := c :: !x; c + t)\ xs\ 0$$
$$\{res. x \mapsto (xs ++ a) * [res = \text{sum } xs]\}$$

How do we prove it automatically?

The traditional approach

- Parameterise specification of foldr over invariants/properties

Some clients may want to operate
only on certain kinds of lists

f must preserve the invariant

$$\forall P, Inv, f, xs, l. \left\{ \begin{array}{l} (\forall x, a', ys. \{P\ x * Inv\ ys\ a'\} \rightarrow f(x, a') \{r. Inv\ (x::ys)\ r\}) \\ * isList\ l\ xs * all\ P\ xs * Inv\ []\ a \end{array} \right\}$$

$foldr\ f\ a\ l$

$\{r. isList\ l\ xs * Inv\ xs\ r\}$

(Separation logic) property
relating *suffix* of input list
traversed to result

The traditional approach

- Parameterise specification of foldr over invariants/properties
- Automation is difficult
 - How to infer invariants/properties to be supplied at call sites?
 - How to infer specification? Clients require different parameterisations

Staged logic

- Natively represent effectful behavior in the logic
- The proof can then be done directly by induction
 - Enabling existing techniques for automated inductive proof [Sun 24]

The rest of this talk

- What are the primitives we need? How do proofs work? (Part I)
- How do we mechanise the proof steps in Coq? (Part II)

Syntax of staged logic

Program states Retain ordering of
 unknown calls/effects

$\varphi ::= \text{req } \sigma \mid \text{ens}[r] \sigma \mid f(v, r) \mid \varphi; \varphi \mid \varphi \vee \varphi \mid \exists x. \varphi \mid \forall x. \varphi$

Assertions Program constructs Logical connectives

Syntax of staged logic

$$\varphi ::= \mathbf{req} \sigma \mid \mathbf{ens}[r] \sigma \mid f(v, r) \mid \varphi; \varphi \mid \varphi \vee \varphi \mid \exists x. \varphi \mid \forall x. \varphi$$

$$\{P\} e \{r. Q\} \equiv e ::: \mathbf{req} P; \mathbf{ens}[r] Q$$

$$e; f(a, r) ::: \mathbf{req} P; \mathbf{ens}[r] Q; f(a, r)$$

foldr

```
let foldr f init xs =  
  match xs with  
  | [] => init  
  | h :: t =>  
    f h (foldr f init t)
```

$$\begin{aligned} \text{foldr}(f, \text{init}, xs, res) = & \\ & \text{ens}[res] \, xs=[] \wedge res=\text{init} \\ & \vee \exists h, t . \text{ens } xs=h::t; \\ & \quad \exists r . \text{foldr}(f, \text{init}, t, r); f(h, r, res) \end{aligned}$$

Giving a specification: entailment

$foldr(g, init, xs, res)$

$\sqsubseteq \mathbf{req} \ x \mapsto a; \mathbf{ens}[res] \ x \mapsto (xs ++ a) * [res = sum\ xs]$

where $(\mathbf{fun} \ c \ t \rightarrow x := c :: !x; c + t) :: g$

A proof strategy

1. Choose argument to perform induction on
2. Unfold non-recursive predicates
3. Rewrite using lemmas/induction hypothesis
4. Normalize, reaching the form

$$(\text{req } \sigma; \text{ens } \sigma; f(a, r);)^* \text{req } \sigma; \text{ens}[r] \sigma$$

5. Dispatch proof obligations using entailment rules

The proof, very briefly

By well-founded induction on xs

$foldr(g, init, xs, res)$

$\sqsubseteq \mathbf{req} \ x \mapsto a; \mathbf{ens} [res] \ x \mapsto (xs ++ a) * [res = sum \ xs]$

The proof, very briefly

Unfold foldr (and focus on the recursive case)

$foldr(g, init, xs, res)$

$\sqsubseteq \mathbf{req} \ x \mapsto a; \mathbf{ens} [res] \ x \mapsto (xs ++ a) * [res = sum \ xs]$

The proof, very briefly

Unfold foldr (and focus on the recursive case)

$$\begin{aligned} & \mathbf{ens} \ (xs = h :: t); \mathit{foldr}(f, \mathit{init}, t, r); f(h, r, res) \\ \sqsubseteq \mathbf{req} \ x \mapsto a; \mathbf{ens}[res] \ x \mapsto (xs ++ a) * [res = \mathit{sum} \ xs] \end{aligned}$$

The proof, very briefly

Unfold foldr (and focus on the recursive case)

ens ($xs = h :: t$);

$foldr(f, init, t, r)$;

$f(h, r, res)$

\sqsubseteq **req** $x \mapsto a$; **ens** $[res]$ $x \mapsto (xs ++ a) * [res = sum\ xs]$

The proof, very briefly

Rewrite using the IH

$$t \leq xs$$

```
ens (xs = h :: t);  
foldr(f, init, t, r);  
f(h, r, res)
```

```
⊢ req x ↦ a; ens[res] x ↦ (xs ++ a) * [res = sum xs]
```

The proof, very briefly

Rewrite using the IH

ens ($xs = h :: t$);

$\forall a_1, \mathbf{req} \ x \mapsto a_1; \mathbf{ens} \ (x \mapsto (t ++ a_1) * [r = \mathit{sum} \ t]);$

$f(h, r, res)$

$\sqsubseteq \mathbf{req} \ x \mapsto a; \mathbf{ens} [res] \ x \mapsto (xs ++ a) * [res = \mathit{sum} \ xs]$

The proof, very briefly

Unfold f

$\mathbf{ens} (xs = h :: t);$

$\forall a_1, \mathbf{req} x \mapsto a_1; \mathbf{ens} (x \mapsto (t ++ a_1) * [r = \mathit{sum} t]);$

$f(h, r, res)$

$\sqsubseteq \mathbf{req} x \mapsto a; \mathbf{ens}[res] x \mapsto (xs ++ a) * [res = \mathit{sum} xs]$

The proof, very briefly

Unfold f

$\mathbf{ens} (xs = h :: t);$

$\forall a_1, \mathbf{req} x \mapsto a_1; \mathbf{ens} (x \mapsto (t ++ a_1) * [r = \mathit{sum} t]);$

$\forall z, \mathbf{req} x \mapsto z; \mathbf{ens}[res] x \mapsto (h :: z) * [res = h + r]$

$\sqsubseteq \mathbf{req} x \mapsto a; \mathbf{ens}[res] x \mapsto (xs ++ a) * [res = \mathit{sum} xs]$

The proof, very briefly

Normalise

$$\frac{H_A * H_1 \vdash H_2 * H_F}{\text{ens } H_1; \text{req } H_2 \sqsubseteq \text{req } H_A; \text{ens } H_F} \text{ NormEnsReq}$$

$\text{ens } (xs = h :: t);$

$\forall a_1, \text{req } x \mapsto a_1; \text{ens } (x \mapsto (t ++ a_1) * [r = \text{sum } t]);$

$\forall z, \text{req } x \mapsto z; \text{ens } [res] x \mapsto (h :: z) * [res = h + r]$

$\sqsubseteq \text{req } x \mapsto a; \text{ens } [res] x \mapsto (xs ++ a) * [res = \text{sum } xs]$

The proof, very briefly

Normalise

$$(z = (t ++ a_1)) * (x \mapsto (t ++ a_1) * [r = \text{sum } t]) \vdash (x \mapsto z) * ([r = \text{sum } t])$$

ens ($xs = h :: t$);

$\forall a_1, \text{req } x \mapsto a_1; \text{ens } (x \mapsto (t ++ a_1) * [r = \text{sum } t]);$

$\forall z, \text{req } x \mapsto z; \text{ens}[res] \ x \mapsto (h :: z) * [res = h + r]$

$\sqsubseteq \text{req } x \mapsto a; \text{ens}[res] \ x \mapsto (xs ++ a) * [res = \text{sum } xs]$

The proof, very briefly

Normalise

$$(z = (t ++ a_1)) * (x \mapsto (t ++ a_1) * [r = \text{sum } t]) \vdash (x \mapsto z) * ([r = \text{sum } t])$$

ens ($xs = h :: t$);

$\forall a_1, \text{req } x \mapsto a_1; \text{ens } (x \mapsto (t ++ a_1) * [r = \text{sum } t]);$

$\forall z, \text{req } x \mapsto z; \text{ens}[res] x \mapsto (h :: z) * [res = h + r]$

$\sqsubseteq \text{req } x \mapsto a; \text{ens}[res] x \mapsto (xs ++ a) * [res = \text{sum } xs]$

The proof, very briefly

Normalise

$$(z = (t ++ a_1)) * (x \mapsto (t ++ a_1) * [r = \text{sum } t]) \vdash (x \mapsto z) * ([r = \text{sum } t])$$

$\forall a_1, \mathbf{req } x \mapsto a_1;$

$\mathbf{ens}[res] x \mapsto (h :: (t ++ a_1)) *$

$[res = h + r \wedge r = \text{sum } t \wedge xs = h :: t]$

$\sqsubseteq \mathbf{req } x \mapsto a; \mathbf{ens}[res] x \mapsto (xs ++ a) * [res = \text{sum } xs]$

The proof, very briefly

We have reached normal form

$\forall a_1, \mathbf{req} \ x \mapsto a_1;$

$\mathbf{ens}[res] \ x \mapsto (h :: (t ++ a_1)) *$

$[res = h + r \wedge r = \mathit{sum} \ t \wedge xs = h :: t]$

$\sqsubseteq \mathbf{req} \ x \mapsto a; \mathbf{ens}[res] \ x \mapsto (xs ++ a) * [res = \mathit{sum} \ xs]$

The proof, very briefly

Contravariance of req

$$\frac{H_2 \vdash H_1}{\text{req } H_1 \sqsubseteq \text{req } H_2} \text{EntailsReq}$$

$\forall a_1, \text{req } x \mapsto a_1;$

$\text{ens}[res] x \mapsto (h :: (t ++ a_1)) *$

$[res = h + r \wedge r = \text{sum } t \wedge xs = h :: t]$

$\sqsubseteq \text{req } x \mapsto a; \text{ens}[res] x \mapsto (xs ++ a) * [res = \text{sum } xs]$

The proof, very briefly

Contravariance of req

$$x \mapsto a \vdash x \mapsto a$$

$\forall a_1, \text{req } x \mapsto a_1;$

$\text{ens}[res] x \mapsto (h :: (t ++ a_1)) *$

$$[res = h + r \wedge r = \text{sum } t \wedge xs = h :: t]$$

$\sqsubseteq \text{req } x \mapsto a; \text{ens}[res] x \mapsto (xs ++ a) * [res = \text{sum } xs]$

The proof, very briefly

Contravariance of req

$$\begin{aligned} & \mathbf{ens}[res] \, x \mapsto (h :: (t ++ a)) * \\ & \quad [res = h + r \wedge r = \mathit{sum} \, t \wedge xs = h :: t] \\ & \sqsubseteq \mathbf{ens}[res] \, x \mapsto (xs ++ a) * [res = \mathit{sum} \, xs] \end{aligned}$$

The proof, very briefly

Covariance of ens

$$\frac{Q_1 \vdash Q_2}{\mathbf{ens} Q_1 \sqsubseteq \mathbf{ens} Q_2} \text{EntailsEns}$$

$$\begin{aligned} \mathbf{ens}[res] x \mapsto (h :: (t ++ a)) * \\ [res = h + r \wedge r = \mathit{sum} t \wedge xs = h :: t] \\ \sqsubseteq \mathbf{ens}[res] x \mapsto (xs ++ a) * [res = \mathit{sum} xs] \end{aligned}$$

The proof, very briefly

Separation logic entailment

$$\begin{aligned} x &\mapsto (h :: (t ++ a)) * \\ &\quad [res = h + r \wedge r = \text{sum } t \wedge xs = h :: t] \\ \vdash x &\mapsto (xs ++ a) * [res = \text{sum } xs] \end{aligned}$$

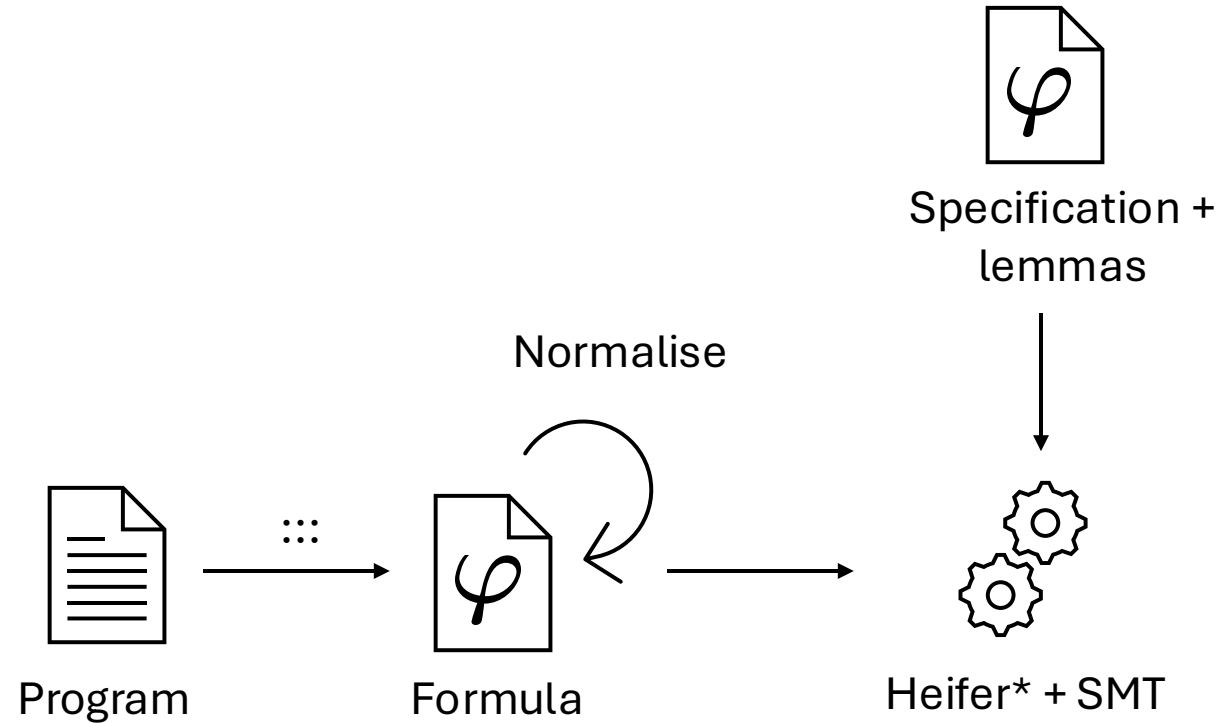
The proof, very briefly

SMT (and some properties of append and cons)

$$res = h + (sum\ t) \wedge xs = h :: t$$

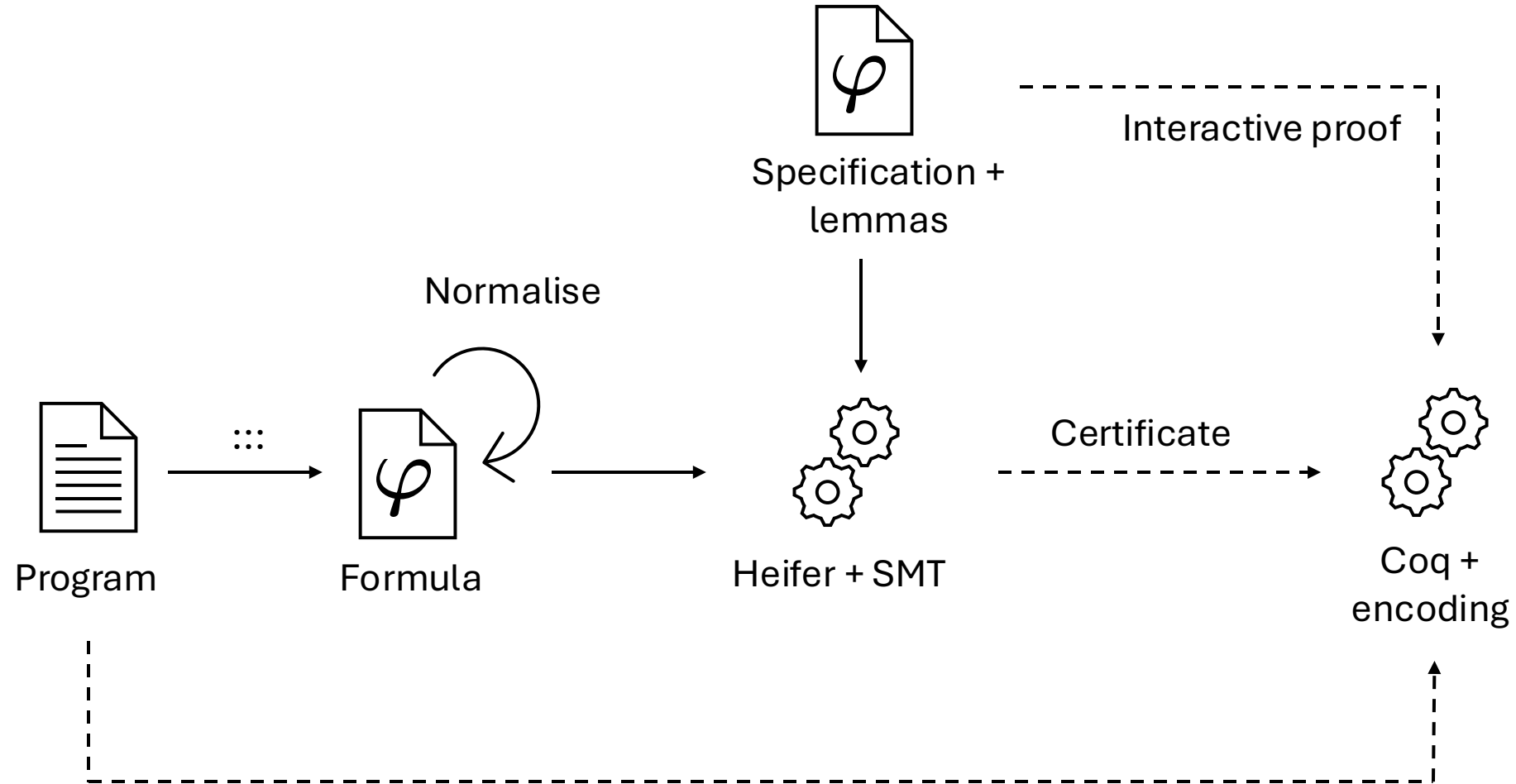
$$\Rightarrow h :: (t ++ a) = xs ++ a \wedge res = sum\ xs$$

The workflow



*Higher-order Effectful Imperative Function Entailments & Reasoning

The workflow we would like



What we would like

▼Goal (1)

How to encode $\mathcal{E} \vdash \varphi \sqsubseteq \varphi$?

```
xs : list val
IH : forall y : list val,
    list_sub y xs →
    forall res0 : val,
    foldr_env ⊢ "foldr"$(vtup (vstr "f") (vtup (vint 0) (vlist y))), res0)
    ⊑ ∀ (x : loc) (a : list val),
        req (x ↗ vlist a)
            (ens_ (x ↗ vlist (y ++ a) \* \[res0 = vint (sum (to_int_list y))]))
res : val
```

Coq sequent

```
foldr_env
⊢ ∃ (x : int) (l1 : list val),
  ens_ \[xs = vint x :: l1];;
  (∃ r : val,
    ("foldr"$(vtup (vstr "f") (vtup (vint 0) (vlist l1))), r));;
    ("f"$(vtup (vint x) r, res)))
⊑ ∀ (x : loc) (a : list val),
  req (x ↗ vlist a)
      (ens_ (x ↗ vlist (xs ++ a) \* \[res = vint (sum (to_int_list xs))]))
```

Staged logic sequent

Semantics of staged logic

$$\begin{array}{ccc}
 & & h \models H \\
 & & h, v \models Q \\
 \mathcal{E}, h_1, h_2, v \models \varphi & & \\
 \swarrow \quad \uparrow \quad \swarrow & & \\
 & \text{heaps} & \text{result}
 \end{array}$$

$$\begin{array}{l}
 \mathcal{E}, h_1, h_2, v \models \mathbf{req} P \varphi \text{ if} \\
 \forall h_p h_r, (h_p \models P \text{ and } h_1 = h_p \circ h_r) \Rightarrow \mathcal{E}, h_r, h_2, v \models \varphi
 \end{array}$$

$$\begin{array}{l}
 \mathcal{E}, h_1, h_2, v \models \mathbf{ens} Q \text{ if} \\
 \exists h_3, (h_3, v \models Q) \text{ and } h_2 = h_1 \circ h_3
 \end{array}$$

Internalization of the operational
behaviour of heap entailment

Semantics of staged logic

An environment of unknown functions

$$\begin{array}{c} \downarrow \\ \mathcal{E}, h_1, h_2, v \models \varphi \\ \begin{array}{ccc} \swarrow & \uparrow & \swarrow \\ & \text{heaps} & \text{result} \end{array} \end{array}$$

$$\begin{array}{l} \mathcal{E}, h_1, h_2, v \models f(a, r) \text{ if} \\ \mathcal{E}, h_1, h_2, v \models \mathcal{E}[f](a, r) \end{array}$$

Encoding staged logic

An environment of unknown functions

$$\mathcal{E}, h_1, h_2, v \models \varphi$$

↖ ↗ ↖
heaps result

Separation logic: heap \rightarrow **Prop**

Unfortunately, a direct shallow embedding would be impredicative

HOAS encoding, enabling substitution

Definition $\text{ufun} := \text{val} \rightarrow \text{val} \rightarrow \text{phi}$.

Definition $\text{phi} := \text{map var ufun} \rightarrow \text{heap} \rightarrow \text{heap} \rightarrow \text{val} \rightarrow \text{Prop}$.

Encoding staged logic

Separation logic: heap \rightarrow **Prop**

Use a deep embedding and interpretation function

An environment of unknown functions

$$\begin{array}{c} \downarrow \\ \mathcal{E}, h_1, h_2, v \models \varphi \\ \begin{array}{ccc} \nwarrow & \uparrow & \nwarrow \\ & \text{heaps} & \text{result} \end{array} \end{array}$$

Inductive phi : **Type** :=
| req : hprop \rightarrow phi \rightarrow phi
| ens : (val \rightarrow hprop) \rightarrow phi
| seq : phi \rightarrow phi \rightarrow phi
| unk : var \rightarrow val \rightarrow val \rightarrow phi
...

Definition ufun := val \rightarrow val \rightarrow phi.

Inductive satisfies :
map var ufun \rightarrow
heap \rightarrow heap \rightarrow val \rightarrow phi \rightarrow **Prop** := ...

Encoding staged logic

Entailment: $\varphi \sqsubseteq \varphi$

We use a semantic definition:

Definition entails (f1 f2:phi) : **Prop** :=
forall env h1 h2 R,
satisfies env h1 h2 R f1 ->
satisfies env h1 h2 R f2.

Lemmas about entailment can be stated and proved directly.

$$\frac{H_1 \vdash H_2}{\mathbf{ens} H_1 \sqsubseteq \mathbf{ens} H_2} \text{EntailsEns}$$

Lemma entails_ens : forall H1 H2,
H1 ==> H2 ->
entails (ens H1) (ens H2).

Encoding staged logic

Entailment sequent: $\mathcal{E} \vdash \varphi \sqsubseteq \varphi$

Parameterised over the environment, supporting rules such as:

$$\frac{F = \mathcal{E}(f)}{\mathcal{E} \vdash f(a, r) \sqsubseteq F(a, r)} \text{EntailsUnfold}$$

What we would like

How to encode $\mathcal{E} \vdash \varphi \sqsubseteq \varphi$? 

▼Goal (1)

```
xs : list val
IH : forall y : list val,
    list_sub y xs →
    forall res0 : val,
    foldr_env ⊢ "foldr"$(vtup (vstr "f") (vtup (vint 0) (vlist y)), res0)
    ⊑ ∀ (x0 : loc) (a : list val),
        req (x0 ↗ vlist a)
        (ens_ (x0 ↗ vlist (y ++ a) \* \[res0 = vint (sum (to_int_list y))]))

res : val
x : int
l1 : list val
H : xs = vint x :: l1
r : val
```

Rewrite

```
foldr_env
⊢ ("foldr"$(vtup (vstr "f") (vtup (vint 0) (vlist l1)), r));;
```

Unfold

```
⊢ ("f"$(vtup (vint x) r, res))
⊑ ∀ (x0 : loc) (a : list val),
    req (x0 ↗ vlist a)
    (ens_ (x0 ↗ vlist (xs ++ a) \* \[res = vint (sum (to_int_list xs))]))
```

Rewriting

- We use Coq's *setoid rewriting*, with entails as the “equivalence” relation
- entails must be shown to be *proper* in both arguments

$$\frac{\varphi_1 \sqsubseteq \varphi_2 \quad \varphi_1 \sqsupseteq \varphi_3 \quad \varphi_2 \sqsubseteq \varphi_4 \quad \varphi_3 \sqsubseteq \varphi_4}{\text{ProperEntails}}$$

Contravariance

Covariance

Rewriting

- This can be specified by providing the following typeclass instance

```
#[global]
Instance Proper_entails : Proper
  (flip entails ===> entails ===> impl)
  entails.
```

What we would like

Rewriting 

$$\frac{\mathcal{E} \vdash \mathbf{ens} \sigma; \varphi_1 \sqsubseteq \varphi_2}{\mathcal{E} \vdash \varphi_1 \sqsubseteq \mathbf{req} \sigma; \varphi_2} \text{EntailsReqR}$$

```
foldr_env
⊢ req (x0 ↗ vlist a)
    (ens_ (x0 ↗ vlist (l1 ++ a) \* \[r = vint (sum (to_int_list l1))]);;
    g (vtup (vint x) r) res)
⊑ req (x0 ↗ vlist a)
    (ens_ (x0 ↗ vlist (xs ++ a) \* \[res = vint (sum (to_int_list xs))]))
```

↑
Can be introduced into
the “spatial context”

What we would like

The spatial context, or the “current state” in symbolic-execution style verifiers


```
foldr_env
⊢ ens_ (x0 ↗ vlist a) ;;
  req (x0 ↗ vlist a)
    (ens_ (x0 ↗ vlist (l1 ++ a) \* \[r = vint (sum (to_int_list l1))]) ;;
      g (vtup (vint x) r) res)
⊢ ens_ (x0 ↗ vlist (xs ++ a) \* \[res = vint (sum (to_int_list xs))])
```

The precondition of the next “call”,
discharged via biabduction

The postcondition of the next “call”

The final proof obligation



What we would like

“Symbolic execution” using biabduction 

foldr_env

```
⊢ ens_ (x0 ~> vlist (vint x :: l1 ++ a) \* \[res = vint (x + sum (to_int_list l1))])  
⊢ ens_ (x0 ~> vlist (xs ++ a) \* \[res = vint (sum (to_int_list xs))])
```

What we would like

```
x0  vlist (vint x :: l1 ++ a) \* \[res = vint (x + sum (to_int_list l1))\] ==>  
x0  vlist ((vint x :: l1) ++ a) \* \[res = vint (sum (to_int_list (vint x :: l1)))\]
```


What we would like

```
res = vint (x + sum (to_int_list l1)) → res = vint (sum (to_int_list (vint x :: l1)))
```

What we would like

No more goals

The mechanisation at a glance

- Other things formalised:
 - Programs, big-semantics
 - $:::$ (pairs), (history) triples
 - Soundness
- 4700 LoC, on top of [Charguéraud 20]

Takeaways

<https://github.com/dariusf/staged>

- An alternative program logic formulation
 - New primitives; no wands, weakest preconditions, or step-indexing
 - Higher-order + effects
- Other views
 - Refinement between abstract programs
 - Triples with syntactic reasoning
 - Manipulating verification conditions directly
- Future work
 - Automation to support certification
 - Other applications of staged logic [Song 24]

Comparison with CFML

- A characteristic formula is a relation between precondition and postcondition, i.e. $cf : \text{expr} \rightarrow (\text{assertion} \rightarrow \text{assertion} \rightarrow \text{Prop})$
- A staged formula is a syntactic entity whose semantics relates pre- and post-*states*
 - This allows more kinds of syntactic reasoning, e.g. mentioning unknown functions

$$\mathcal{E}, h_1, h_2, v \models \varphi$$

Biabduction

$$H_A * H_1 \vdash H_2 * H_F \quad [\text{Calcagno 09}]$$

Deeply embedded, for induction over derivations

```
Inductive biab : hprop -> hprop -> hprop -> hprop -> Prop :=  
| b_base_empty : forall Hf,  
  biab \[] Hf \[] Hf  
  
| b_pts_match : forall a b H1 H2 Ha Hf x,  
  biab Ha H1 H2 Hf ->  
  biab (\[a=b] \* Ha) (x~~>a \* H1) (x~~>b \* H2) Hf  
...  

```

```
Lemma biab_sound : forall Ha H1 H2 Hf,  
  biab Ha H1 H2 Hf ->  
  Ha \* H1 ==> H2 \* Hf.
```

Why mechanise an automated verifier?

- Check claims in paper
- Clarify ideas: find the simplest version of each concept
- Communicate: other people can try the logic and build intuition
- Certification: validate implementation, not just ideas
- Support new work: work manually on harder proofs, broaden fragment that can be fully automated