Integrating Iris into the Verified Software Toolchain, and vice versa

William Mansky, University of Illinois Chicago Iris Workshop, May 23rd, 2023

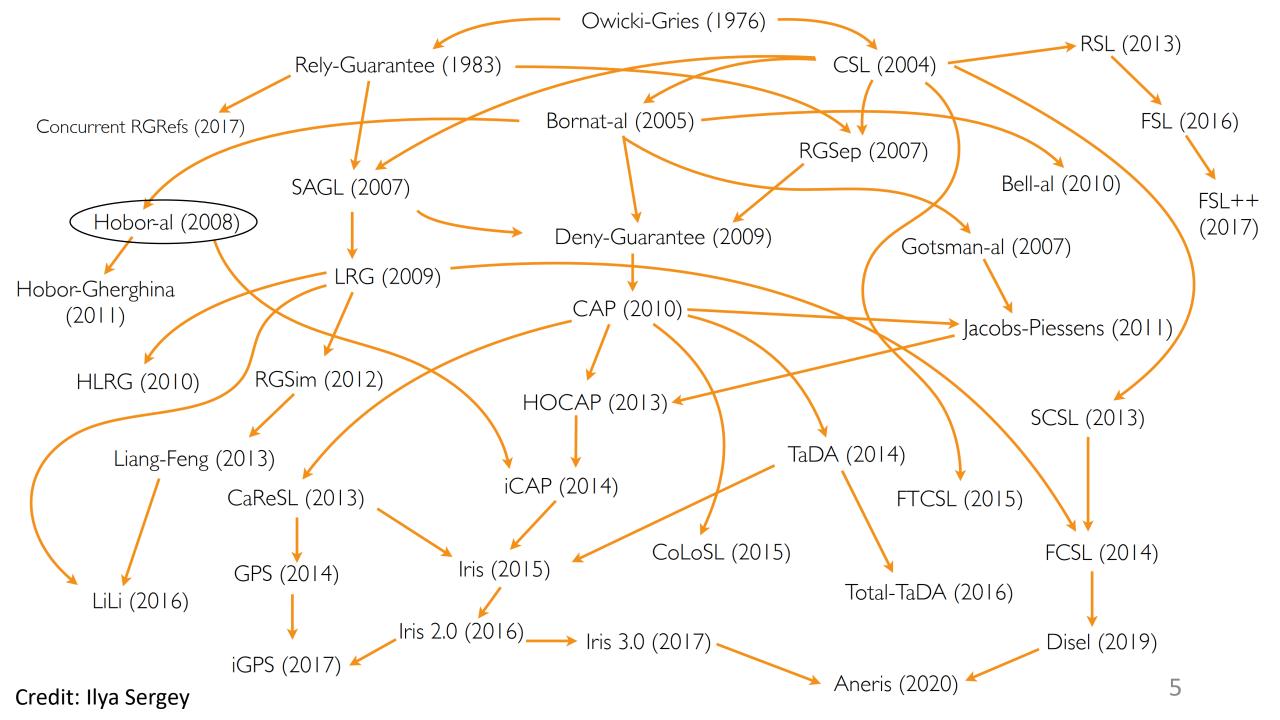


The Verified Software Toolchain (VST)

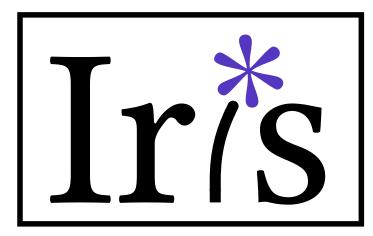
Concurrent separation logic verifier in Coq

Verified Software Toolchain

- Higher-order, step-indexed, symbolic execution + entailment solving
- Specialized to C, connected to CompCert
 - End-to-end soundness theorem



Iris and VST

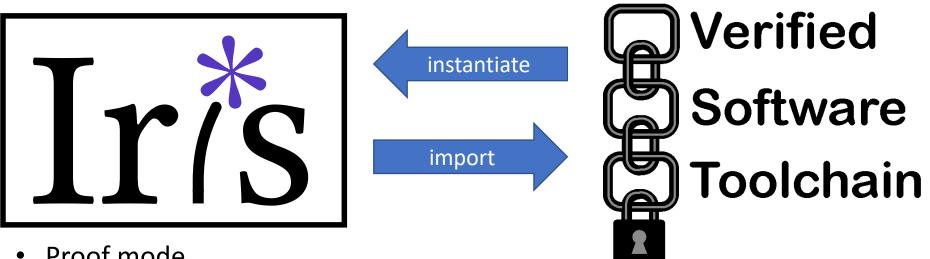


- Proof mode
- Custom ghost state
- Invariants
- Logical atomicity
- ...



- C isn't garbage-collected, so logic shouldn't be affine
- Ownership can't be "just ghost state": it's translated to CompCert permissions and used for adequacy

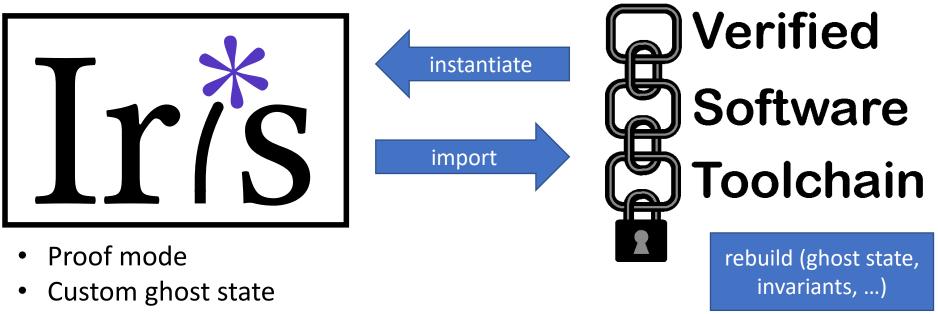
Iris in VST



- Proof mode
- Custom ghost state
- Invariants
- Logical atomicity
- ...

• Keep VST's foundations the same, import or reconstruct the features we want

Iris in VST



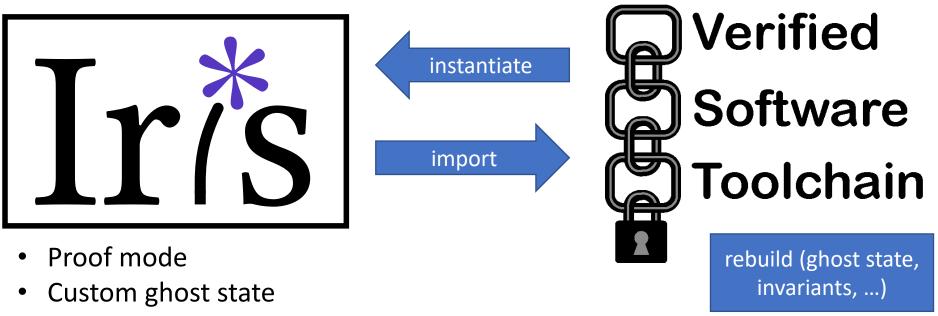
- Invariants •
- Logical atomicity
- ...
- Keep VST's foundations the same the core of the model and soundness proof, import or reconstruct the features we want

VST on Iris



 Replace VST's foundations with Iris, rebuild the rest of VST on top, get Iris features for free

Iris in VST



- Invariants •
- Logical atomicity
- ...
- Keep VST's foundations the same the core of the model and soundness proof, import or reconstruct the features we want

VST + MoSeL

Ir/s

Verified Software Toolchain

Get IPM by instantiating the BI interface with VST's logic Non-obvious parts of BI:

- Step-indexing and arpropto but they're exactly the same in VST as in Iris
- Persistence (⊡) modality
 - Default definition: $(\bigcirc P)(x)$ when $(P)(\operatorname{core} x)$
 - VST has core too, but core is always emp!
 - Simple instantiation: $(\bigcirc P)(x)$ when $(P \land emp)(x)$

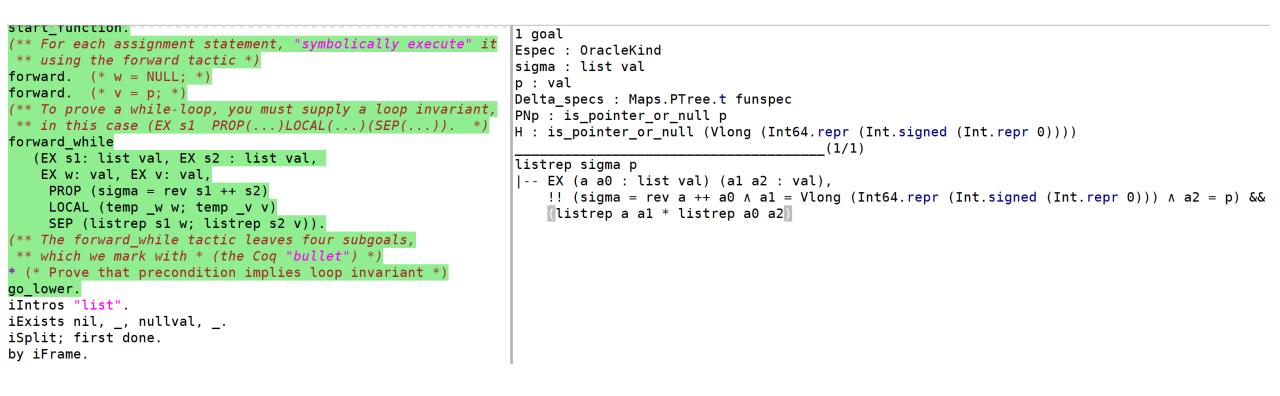
MoSeL: A General, Extensible Modal Framework for Interactive Proofs in Separation Logic, Krebbers et al., ICFP 2018

VST + MoSeL



Verified Software Toolchain

Immediately get MoSeL tactics for VST's separation logic



VST + MoSeL



Verified Software Toolchain

Immediately get MoSeL tactics for VST's separation logic

start function. 1 goal (** For each assignment statement, "symbolically execute" it Espec : OracleKind ** using the forward tactic *) sigma : list val forward. (* w = NULL; *)p : val forward. (* v = p; *)Delta specs : Maps.PTree.t funspec (** To prove a while-loop, you must supply a loop invariant, PNp : is pointer or null p ** in this case (EX s1 PROP(...)LOCAL(...)(SEP(...)). *) H : is pointer or null (Vlong (Int64.repr (Int.signed (Int.repr 0)))) forward while (1/1)(EX s1: list val, EX s2 : list val, "list" : listrep sigma p EX w: val, EX v: val, PROP (sigma = rev s1 ++ s2) EX (a a0 : list val) (a1 a2 : val), LOCAL (temp _w w; temp _v v) !! (sigma = rev a ++ a0 \wedge a1 = Vlong (Int64.repr (Int.signed (Int.repr 0))) \wedge a2 = p) && SEP (listrep s1 w; listrep s2 v)). (listrep a al * listrep a0 a2) (** The forward while tactic leaves four subgoals, ** which we mark with * (the Cog "bullet") *) * (* Prove that precondition implies loop invariant * go lower. iIntros "list". iExists nil, _, nullval, . iSplit; first done. by iFrame.

VST + MoSel



Verified Software Toolchain

Immediately get MoSeL tactics for VST's separation logic

- Don't get wp tactics, but VST has its own (forward)
- Don't yet get tactics for invariants, updates, atomic updates, etc. those have their own classes to instantiate
- But first, we need those things to exist in VST!

MoSeL: A General, Extensible Modal Framework for Interactive Proofs in Separation Logic, Krebbers et al., ICFP 2018

Ghost State in VST



- Model of Iris: $M \triangleq \prod_{i \in I} \mathbb{N} \rightharpoonup M_i$, where cameras M_i may include predicates
- Model of VST ("rmap"): $R \triangleq loc \rightarrow res$, where res may include predicates ("predicates in the heap")
- New model of VST: R * M
 - In practice, we don't get HO ghost state, just agreement

Ghost State Updates in VST



- New model of VST: *R* * *M* (approximately)
- Now we can define ⇒, and add updates between steps in our semantics
 - And instantiate BUpd class, use iMod and iModIntro
- Hoare rules are unchanged, since program ops ignore ghost state
- Adequacy proof basically unchanged
- Used for simple double-increment example, external state reasoning

Verifying an HTTP Key-Value Server with Interaction Trees and VST, Zhang et al., ITP 2021

Invariants in VST

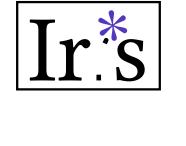


- Invariants can be built out of ghost state
- In Iris: $W \triangleq \bullet(\bullet I) * \overset{\bullet}{} (\bullet I(i) * \operatorname{dis}(i)) \lor \operatorname{en}(i)$ where *I* is a map from names to assertions
- This is exactly what we can't do with our restricted HO ghost state!
- Refactored construction:

$$W \triangleq \bullet G * \underset{i \in \text{dom}(G)}{} \text{ag}_{G(i)}I(i) * ((\triangleright I(i) * \text{dis}(i)) \lor \text{en}(i))$$

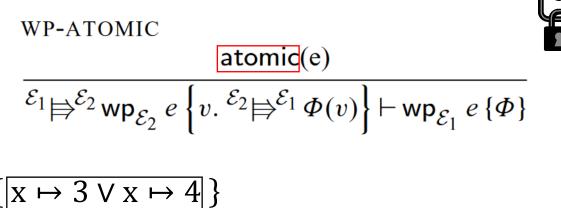
• Satisfies all the same proof rules, and we can build namespaces, instantiate proofmode classes for invariants, etc. on top of it

Fancy Updates in VST



Verified Software Toolchain

 $\begin{array}{l} & \text{WP-VUP} \\ & \rightleftharpoons_{\mathcal{E}} \operatorname{wp}_{\mathcal{E}} e\left\{ v. \rightleftharpoons_{\mathcal{E}} \Phi(v) \right\} \vdash \operatorname{wp}_{\mathcal{E}} e\left\{ \Phi \right\} \end{array}$



$$\{ \begin{array}{c} x \mapsto 3 \lor x \mapsto 4 \\ x \leftarrow 3 \mid \mid x \leftarrow 4 \\ \{ \begin{array}{c} x \mapsto 3 \lor x \mapsto 4 \end{array} \} \end{array}$$

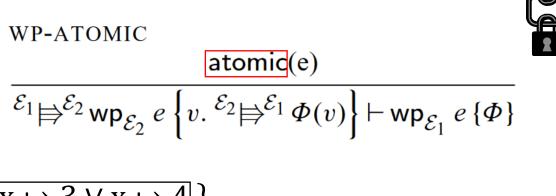
• In C, this is undefined behavior!

Fancy Updates in VST



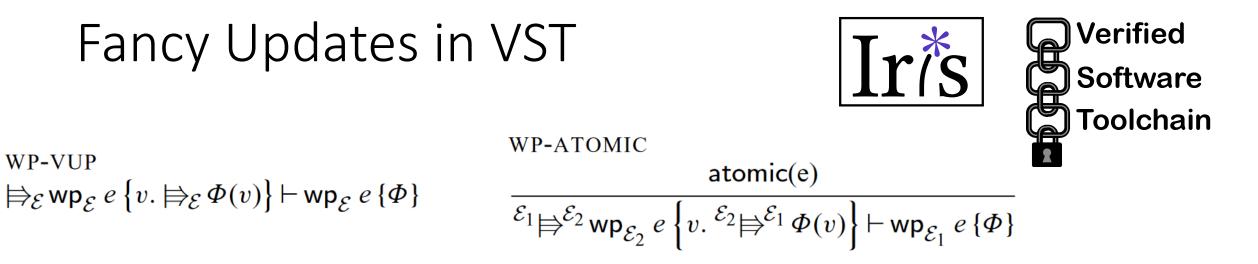
Verified Software Toolchain

WP-VUP $\models_{\mathcal{E}} wp_{\mathcal{E}} e\left\{v. \models_{\mathcal{E}} \Phi(v)\right\} \vdash wp_{\mathcal{E}} e\left\{\Phi\right\}$



$$\{ \begin{array}{c} x \mapsto 3 \lor x \mapsto 4 \\ x \leftarrow 3 \mid \mid x \leftarrow 4 \\ \{ x \mapsto 3 \lor x \mapsto 4 \} \end{array}$$

- In C, this is undefined behavior!
- We set atomic to mean concurrency-atomic: lock acquire/release, atomic_load/store, etc., and nothing else



- We set atomic to mean concurrency-atomic: lock acquire/release, atomic_load/store, etc.
- Unlike basic updates, this changes the semantics: "real" resources can change hands between steps
- For concurrent soundness, have to prove race-freedom, which seems true but not obvious

Persistence in VST



Verified Software Toolchain

- In Iris: invariants are *persistent*, can freely be automatically duplicated and passed between threads
- In VST: we defined $\bigcirc P$ to only hold on emp!
- Invariants really need to be affine too, but in VST nothing is affine!
- Step 1: weaken the core axiom
- Step 2: make the logic *semi-linear*

Persistence in VST

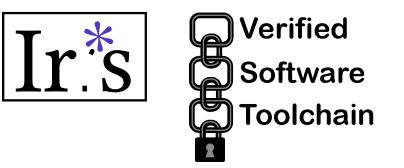


Step 1: weaken the core axiom

VST: $a \leq b \rightarrow \operatorname{core}(b) = \operatorname{core}(a)$ Iris: $a \leq b \rightarrow \operatorname{core}(b) \leq \operatorname{core}(a)$

- In Iris, the core of (N, +) can tell us "the value is at least n"; in VST, it can only tell us "the value is at least 0"
- Simple solution: weaken VST's core axiom
 - Heap resources still have trivial cores, but ghost state doesn't have to
 - Now we can define useful persistence
 - And all the existing proofs still work

Persistence in VST



Step 2: make the logic *semi-linear*

- ORA idea from MoSeL: equip algebras with an *extension order* ⊑ describing which resources can be thrown away
- Define predicates to be closed under ⊑
- Surprisingly, VST's model also has a slot for this order! Included in 2009, never mentioned in a paper or instantiated nontrivially
- We choose: $(r,m) \sqsubseteq (r',m') \triangleq r' = r \land m \preccurlyeq m'$
- Now all ghost state is affine, and all ghost state cores (including invariants) are intuitionistic!

Using Iris in VST



Verified Software Toolchain

- We now have custom ghost state, invariants, updates, and all the relevant Iris tactics in VST
- Can import definitions like logical atomicity directly

- Intros i il keys; forward. forward.	keys : list Z
rewrite -> sub_repr, and_repr; simpl.	H5 : il `mod` size = (i + hash k) `mod` size
rewrite -> Zland_two_p with (n := 14) by lia.	H6 : $0 \le i < size$
replace (2 ^ 14) with size by (setoid_rewrite (proj2_sig has_size); auto).	H7 : Zlength keys = size
exploit (Z_mod_lt il size); [lia intro Hil].	H8 : Forall (λ z : Z, (z ≠ 0 Λ z ≠ k)%type) (sublist 0 i (rebase keys (hash k)))
assert_PROP (Zlength entries = size) as Hentries by entailer!.	Hi1 : 0 ≤ i1 `mod` size < size
assert (0 <= i1 mod size < Zlength entries) as Hil' by lia.	Hentries : Zlength entries = size
match goal with H : Forall => pose proof (Forall_Znth Hi1' H) as Hptr end.	Hil' : 0 ≤ il `mod` size < Zlength entries
destruct (Znth (il mod size) entries) as (pki, pvi) eqn: Hpi; destruct Hptr.	pki, pvi : val
forward; setoid_rewrite Hpi.	Hpi : Znth (il `mod` size) entries = (pki, pvi)
{ entailer!. }	H9 : isptr pki
assert (Zlength (rebase keys (hash k)) = size) as Hrebase.	H10 : isptr pvi
<pre>{ rewrite Zlength_rebase; replace (Zlength keys) with size; auto; apply hash_range. }</pre>	MORE_COMMANDS := abbreviate : statement
forward_call atomic_load_int (pki, top, empty,	Hrebase : Zlength (rebase keys (hash k)) = size
fun v : Z => AS * ghost_snap v (Znth (i1 mod size) lg)).	(1/1)
<pre>{ rewrite !sepcon_assoc; apply sepcon_derives; [cancel].</pre>	semax Delta
iIntros ">AS".	(PROP ()
iDestruct ("AS") as (HT) "[hashtable Hclose]"; simpl.	LOCAL (temp _i pki; temp _idx (vint (i1 `mod` size)); lvar _ref tint v_ref;
iDestruct "hashtable" as (T) "((% & excl) & entries)".	temp _key (vint k); temp _value (vint v); gvars gv)
rewrite -> @iter_sepcon_Znth' with (d := Inhabitant_Z) (i := i1 mod size) by	<pre>SEP (AS; data_at Tsh tint (vint 0) v_ref; data_at sh (tarray tentry size) entries (gv _m_entries);</pre>
(try apply Cveric; rewrite Zlength_upto Z2Nat.id; lia).	iter_sepcon
erewrite Znth_upto by (rewrite -> ?Zlength_upto, Z2Nat.id; lia).	$(\lambda$ i0 : Z, ghost_snap (Znth ((i0 + hash k) `mod` size) keys) (Znth ((i0 + hash k) `mod` size) lg
unfold hashtable_entry at 1.	(upto (Z.to_nat i)))) ((_t'2 = _atom_load(_i);
rewrite Hpi.	$probed_key = t'2;$
destruct (Znth (il mod size) T) as (ki, vi) eqn: HHi.	MORE_COMMANDS) POSTCONDITION

Using Iris in VST



Verified Software Toolchain

- We now have custom ghost state, invariants, updates, and all the relevant Iris tactics in VST
- Can import definitions like logical atomicity directly

```
Intros i il keys; forward. forward.
rewrite -> sub_repr, and_repr; simpl.
rewrite -> Zland two p with (n := 14) by lia.
replace (2 ^ 14) with size by (setoid_rewrite (proj2_sig has_size); auto).
exploit (Z mod lt i1 size); [lia | intro Hi1].
assert_PROP (Zlength entries = size) as Hentries by entailer!.
assert (0 <= i1 mod size < Zlength entries) as Hi1' by lia.
match goal with H : Forall _ _ |- _ => pose proof (Forall_Znth _ _ _ Hi1' H) as Hptr end.
destruct (Znth (i1 mod size) entries) as (pki, pvi) eqn: Hpi; destruct Hptr.
forward; setoid rewrite Hpi.
{ entailer!.
assert (Zlength (rebase keys (hash k)) = size) as Hrebase.
{ rewrite Zlength rebase; replace (Zlength keys) with size; auto; apply hash range.
forward call atomic load int (pki, top, empty,
  fun v : Z \Rightarrow AS * ghost snap v (Znth (i1 mod size) lg)).
{ rewrite !sepcon assoc; apply sepcon derives; [|cancel].
  iIntros ">AS".
  simpl.
  iDestruct ("AS") as (HT) "[hashtable Hclose]"; simpl.
  iDestruct "hashtable" as (T) "((% & excl) & entries)".
  rewrite -> @iter sepcon Znth' with (d := Inhabitant Z) (i := i1 mod size) by
                                                                                              |=\{empty\}=>
    (try apply Cveric; rewrite Zlength upto Z2Nat.id; lia).
  erewrite Znth upto by (rewrite -> ?Zlength upto, Z2Nat.id; lia).
  unfold hashtable entry at 1.
  rewrite Hpi.
```

V 3 TT 1100 3TTC - 3TTC Hentries : Zlength entries = size Hil' : $0 \le i1$ `mod` size < Zlength entries pki, pvi : val Hpi : Znth (il `mod` size) entries = (pki, pvi) H9 : isptr pki H10 : isptr pvi Hrebase : Zlength (rebase keys (hash k)) = size Frame := [data at Tsh tint (vint 0) v ref; data at sh (tarray tentry size) entries (gv m entries); iter sepcon (λ i : Z. ghost snap (Znth ((i + hash k) `mod` size) keys) (Znth ((i + hash k) `mod` size) lg)) (upto (Z.to_nat i))] : list mpred (1/1)"AS" : $\exists x : Z \rightarrow option Z$, hashtable x g lg entries *(hashtable x g lg entries ={empty,top}=* AU << $\exists \exists x0 : Z \rightarrow option Z$, hashtable x0 g lg entries >> d top. empty << $\forall \forall$ _ : (), hashtable (map_upd x0 k v) g lg entries * emp, COMM Q >>) Λ (\forall : (), hashtable (map upd x k v) g lg entries * emp ={empty,top}=* Q) EX (sh0 : share) (v0 : Z), !! (readable share sh0 Λ repable signed v0) && atomic int at sh0 (vint v0) pki * [atomic int at sh0 (vint v0) pki -* (|={empty,top}=> AS * ghost snap v0 (Znth (i1 `mod` size) lg))

Iris in VST: Summary



- All the concurrency features of Iris, in VST
- Foundational changes: ghost state in the model, weaker core axiom, extension order for affine ghost state, fancy updates in the semantics
- Now we can prove atomic specs for concurrent C programs, using VST for C code and switching to Iris tactics for concurrency reasoning
- Could be useful for other non-Iris verifiers that want Iris features
- Can now reconstruct ghost-state-based reasoning in VST, e.g. ReLoC
- Paper on arXiv, opam package coq-vst-iris

Iris in VST: Summary



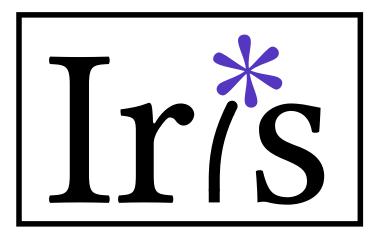
Verified Software Toolchain

• Now we can prove atomic specs for concurrent C programs, is using VST for C code and switching to Iris tactics for concurrency

But:

- Concurrent soundness is still complicated
- We're reconstructing Iris features, and there's always more we might want to reconstruct (transfinite step-indexes, later credits, ...)
- We're working in parallel to RefinedC and the whole Iris ecosystem
- What if VST was built on Iris instead?

Iris and VST



- Proof mode
- Custom ghost state
- Invariants
- Logical atomicity
- ...



- C isn't garbage-collected, so logic shouldn't be affine We can use ORAs!
- Ownership can't be "just ghost state": it's translated to CompCert permissions and used for adequacy
 Need a fancier relationship between physical state and mapsto assertions

VST on Iris



 Replace VST's foundations with Iris, rebuild the rest of VST on top, get Iris features for free

VST on Iris: "juicy" view



- State interpretation: σ where σ is a map from locations to values
- Maps-to: $l \mapsto v$ is defined as $\circ \{[l \coloneqq v]\}$

• $\sigma * l \mapsto v \vdash \sigma(l) = v$

- In VST, these don't coincide!
 - Physical memory (CompCert) maps locations to values + permissions (readable, writable, etc.)
 - Logical memory maps locations to rmap resources + shares
 - Semantics defined in terms of a "juicy mem" that includes both CompCert mem and rmap, plus proof that they are coherent

VST on Iris: "juicy" view



• General *views*: parameterized by a relation *R*, give:

• $a * \circ b \vdash R a b$

• In VST, we can choose $R \triangleq$ **coherent**, and get:

• $m * l \mapsto_{\pi} v \vdash \operatorname{coherent} m l \pi v$

VST on Iris: "juicy" view



39

• General *views*: parameterized by a relation *R*, give:

• $a * \circ b \vdash R a b$

• In VST, we can choose $R \triangleq$ **coherent**, and get:

•
$$m * l \mapsto_{\pi} v \vdash \operatorname{coherent} m l \pi v$$

Old VST: $\forall j. (l \mapsto v) (\operatorname{rmap_of} j) \rightarrow \operatorname{valid_pointer} l (\operatorname{mem_of} j)$ VST on Iris: $\bullet m * l \mapsto v \vdash \ulcorner \operatorname{valid_pointer} l m \urcorner$

VST on Iris: semantics



- Iris: wp $e \{\Phi\}$ when either e is terminated in a state satisfying Φ , or $S(\sigma) \Rightarrow (e, \sigma) \rightarrow (e', \sigma') \Rightarrow S(\sigma') * wp e' \{\Phi\}$
- VST defines *safety* similarly, except that there are two kinds of steps:
 - Core steps are steps by the Clight semantics
 - External calls call arbitrary external functions with provided pre- and postconditions
- Safety was originally defined as a relation on juicy mems, but we can rephrase it inside the logic analogously to wp

VST on Iris: program logic



- Proved exactly the same triples for C statements (mod. Iris notation)
- Proofs are about ½ the size of old versions
 - #1 reduction: reasoning at the logic level instead of unfolding to the model
 - #2 reduction: proof mode tactics

VST on Iris: adequacy



- Still in progress: should be the same paper proof, but in Iris terms
- Aim to prove as much as possible (probably everything!) in the logic instead of unfolding to the model
- VST has complicated armature for lifting CompCert's soundness to concurrency; it should be easier with Iris, but basically the same
- We're long overdue for a better approach to compiler correctness for concurrency! Happy to talk if you have ideas.

VST on Iris: user interface



- Still need to rebuild symbolic execution tactics and automation
- Interaction mode 1: VST + Iris
 - Can do anything we did before in VST in exactly the same way
 - Drop into Iris proof mode as desired for invariants, atomics, etc.
- Interaction mode 2: Iris style
 - Turn Hoare triples into WP format, stay in IPM the whole time
 - Will require retooling VST's automation (forward, etc.) to work on IPM goals
 - More comfortable for Iris people, could adapt Diaframe

Conclusion





- Iris in VST: mostly done
 - Can prove logically atomic specs for C programs using Iris logic and tactics
 - Takes cues from Iris, reuses some of it, rebuilds a lot more
- VST on Iris: looks like it'll work!
 - More expressive ghost state
 - Can incorporate new Iris ideas: transfinite step-indexing, later credits, ...
 - Integrate with other tools? Diaframe, RefinedC, ...
 - What would you do with a CompCert C mode for Iris?

VST on Iris: ownership



- Iris mapsto is simple: $l \mapsto_q v$, where q is a positive fraction
 - Any q is enough to read, 1 is required to write
- VST uses tree shares, with 4 distinct permission levels (corresponding to CompCert permission levels): nonempty, readable, writable, freeable
- Nonempty ownership gives knowledge of the location, but not its value!

```
Inductive shared :=
| YES (dq : dfrac) (rsh : readable_dfrac dq) (v : agree V)
| NO (sh : shareO) (rsh : ¬readable_share' sh).
```

- $l \mapsto_q v$ is $\{[l \coloneqq \text{YES } q _ v]\}$
- Also have $l \mapsto_q \bot$, which is $\{[l \coloneqq \text{NO } q \ _]\}$

VST on Iris: resources



- Model of VST: $R \triangleq loc \rightarrow res$, where res may include predicates ("predicates in the heap")
- res \triangleq VAL $v \mid$ LK $R \mid$ FUN A P Q
- Last two use "predicates in the heap"
 - But in Iris they don't need to! We'll come back to this



- Model of VST: $R \triangleq loc \rightarrow res$, where res may include predicates ("predicates in the heap")
- res \triangleq VAL $v \mid$ LK $R \mid$ FUN A P Q
- We can take the predicates out of the heap, and use ghost state/invariants for them instead



- Model of VST: $R \triangleq loc \rightarrow res$, where res may include predicates ("predicates in the heap")
- res \triangleq VAL $v \mid$ LK \mid FUN A P Q
- We can take the predicates out of the heap, and use ghost state/invariants for them instead
- isLK $l R \triangleq l \mapsto LK * inv R$



• res \triangleq VAL $v \mid$ LK \mid FUN A P Q

Inductive funspec := mk_funspec: typesig -> calling_convention ->
forall A (P: A -> mpred) (Q: A -> mpred), funspec.

- $l \mapsto \text{FUN } A P Q$ asserts that l is a function pointer w/ spec $\forall a: A, \{P \ a\} l \{Q \ a\}$
- We build an OFE for funspec (roughly isomorphic to $\{A \& (A \rightarrow mpred) * (A \rightarrow mpred)\}\)$



• res \triangleq VAL $v \mid$ LK \mid FUN

Inductive funspec := mk_funspec: typesig -> calling_convention ->
forall A (P: A -> mpred) (Q: A -> mpred), funspec.

- $l \mapsto \text{FUN } A P Q$ asserts that l is a function pointer w/ spec $\forall a: A, \{P \ a\} l \{Q \ a\}$
- We build an OFE for funspec (roughly isomorphic to {A & (A → mpred) * (A → mpred)})
- Define isFUN $l f \triangleq l \mapsto FUN * \circ \{[l \coloneqq \rhd f]\}$
 - Analogous to invariant construction