

A Separation Logic Framework for Asynchronous Programming

João Mota - jd.mota@campus.fct.unl.pt

Advisers: António Ravara and Marco Giunti

NOVA LINCS and NOVA School of Science & Technology

Producer + Consumer in JavaScript

```
const queue = new Queue();
```

```
async function producer() {  
  const data = await otherTask1();  
  queue.add(data);  
  queue.unuse();  
}
```

```
async function consumer() {  
  while (queue.inUse()) {  
    const data = await queue.take();  
    await otherTask2(data);  
  }  
}
```

```
queue.use();  
await Promise.all([ producer(), consumer() ]);
```

Producer + Consumer in Java

```
Queue queue = new Queue(); // Locks inside
```

```
Thread producer = new Thread(() -> {  
    Data data = otherTask1();  
    queue.add(data);  
    queue.unuse();  
});
```

```
Thread consumer = new Thread(() -> {  
    while (queue.inUse()) {  
        Data data = queue.take();  
        otherTask2(data);  
    }  
});
```

```
queue.use();  
producer.start(); consumer.start();
```

Problem

- Languages have *different concurrency models*
 - Single-threaded asynchrony (e.g., JavaScript)
 - Multi-threading (e.g., Java)
- How to guarantee **memory safety**, **thread-safety**, and **avoid memory leaks**?
 - No distinction between *thread-local* and *thread-shared* data forces access to shared data to be *synchronized*
 - Forced to choose between a fully *affine* or fully *linear* logic

We need a unified approach

Separation Logic for Async. Prog.

Framework heavily inspired on **Iris** ^[1]. Features:

- Affine and linear resources/invariants
- Thread-local and thread-shared invariants
- Invariants are precisely tracked
- Weakest-precondition tailored for a concurrent language with threads and asynchronous tasks in each thread

1. <https://iris-project.org/> ↩

Affine and linear resources

- **affine** resources can be "dropped" (and be garbage collected)
- **linear** resources must be necessarily used
even without `free()`, these are useful to avoid memory leaks

E.g., queue as a linear resource until it is empty

E.g., a socket as a linear resource until calling `close()`

Thread-local/shared invariants

- **thread-shared** invariants are opened around atomic expressions
- **thread-local** may remain open for several steps until reaching a yield-point (such as the *await* keyword)

E.g., queue as a thread-shared resource in Java

E.g., queue as a thread-local resource in JS

Base Logic

- Non-step-indexed
- We extend algebras with an *affinity predicate* \mathcal{A}
- Connectives are obtained from Iris by replacing \multimap with \sqsubseteq [1]
 - $\sqsubseteq \triangleq \exists c. b \equiv a \cdot b \wedge \mathcal{A}(c)$
 - Precise tracking of linear resources requires *cancelability*

1. Krebbers et al. MoSeL: a general, extensible modal framework for interactive proofs in separation logic. ICFP 2018 ↩

Basic Update Modality

- Issue: $P \vdash \Vdash \epsilon$
- Extend the modality with a *controlled update relation*
 - $\Vdash\{u\}P \triangleq$
 $\lambda a. \forall c. \mathcal{V}(a \cdot c) \Rightarrow \exists b. \mathcal{V}(b \cdot c) \wedge P(b) \wedge (a, b) \in u$
- For a logic parameterized with the *authoritative algebra*, we can define the following relation to preserve linear resources:

$$\frac{\forall f. b_1 \cdot f \equiv a_1 \Rightarrow b_2 \cdot f \sqsubseteq a_2}{\bullet a_1, \circ b_1 \dashrightarrow_{linear} \bullet a_2, \circ b_2}$$

...

Linear Update Example

$$\frac{\forall f. b_1 \cdot f \equiv a_1 \Rightarrow b_2 \cdot f \sqsubseteq a_2}{\bullet a_1, \circ b_1 \dashrightarrow_{linear} \bullet a_2, \circ b_2}$$

$$[\iota_2 \leftarrow v_2] \cdot [\iota_1 \leftarrow v_1] \equiv [\iota_1 \leftarrow v_1, \iota_2 \leftarrow v_2]$$

$$\varepsilon \cdot [\iota_1 \leftarrow v_1] \sqsubseteq [\iota_1 \leftarrow v_1, \iota_2 \leftarrow v_2]$$

$$\bullet [\iota_1 \leftarrow v_1, \iota_2 \leftarrow v_2], \circ [\iota_2 \leftarrow v_2] \dashrightarrow_{linear} \bullet [\iota_1 \leftarrow v_1, \iota_2 \leftarrow v_2], \circ \varepsilon$$

Forgetting about reference ι_2

Linear Update Example

$$\frac{\forall f. b_1 \cdot f \equiv a_1 \Rightarrow b_2 \cdot f \sqsubseteq a_2}{\bullet a_1, \circ b_1 \dashrightarrow_{linear} \bullet a_2, \circ b_2}$$

$$\frac{[\iota_2 \leftarrow v_2] \cdot [\iota_1 \leftarrow v_1] \equiv [\iota_1 \leftarrow v_1, \iota_2 \leftarrow v_2] \quad \varepsilon \cdot [\iota_1 \leftarrow v_1] \sqsubseteq [\iota_1 \leftarrow v_1]}{\bullet [\iota_1 \leftarrow v_1, \iota_2 \leftarrow v_2], \circ [\iota_2 \leftarrow v_2] \dashrightarrow_{linear} \bullet [\iota_1 \leftarrow v_1], \circ \varepsilon}$$

Deallocation of reference ι_2

Invariant Masks

To support both thread-shared and thread-local invariants,
we extend *masks* to be sets of pairs,
with *thread-sharedness* and *invariant identifier*

$$ts \in \mathcal{T} \triangleq \text{ThShared} \mid \text{ThLocal}(t : \mathbb{N})$$

where t is a thread identifier

$$\mathcal{E} \in \text{Mask} \triangleq \mathcal{P}(\mathcal{T} \times \mathbb{N})$$

Fancy Update Modalities

The "normal" fancy update modality is the same, but with a globally parameterized controlled update relation

$$\varepsilon_1 \Vdash \varepsilon_2 P \triangleq \text{wsat} * \boxed{\varepsilon_1}^{En} \multimap \dot{\Vdash}\{u\}(\text{wsat} * \boxed{\varepsilon_2}^{En} * P)$$

The novel "predicated" fancy update modality allows to pick the mask that must be reestablished after the update (if it fulfills $\Phi_{\mathcal{E}}$), enabling thread-local invariants to remain opened for longer

$$\varepsilon_1 \dot{\Vdash}_{\Phi_{\mathcal{E}}} \Phi \triangleq \varepsilon_1 \Vdash_{\emptyset} (\exists \mathcal{E}_2. \ulcorner \Phi_{\mathcal{E}}(\mathcal{E}_2) \urcorner \wedge \boxed{\varepsilon_2}^{En} * \Phi(\mathcal{E}_2))$$

World Satisfaction

- Two maps:
 1. Stores the propositions (persistently)
 2. Tracks the allocated invariants
- Allocated invariants in (2) are composed by *external id*, *internal id* to (1), affinity, thread-sharedness
- Propositions in (1) belong to a logic parameterized with user-given resource algebras, excluding the algebra that encodes (1)
- The indirection avoids circularity in the model and allows storing invariant tokens inside of invariants without step-indexing

World Satisfaction

- Invariants are precisely tracked with some given model of permissions (e.g., fractional or counting)
- As in Iron^[1], we have *opening* and *deallocation* tokens
 - Deallocation tokens for linear invariants cannot be put inside of invariants
 - Both kinds can be used to open invariants

1. Ales Bizjak et al. Iron: managing obligations in higher-order concurrent separation logic. POPL 2019 ↩

World Satisfaction

$$\text{holds}(P, \text{aff}) \triangleq \begin{cases} \text{Emp} \wedge \uparrow P & \text{if } \text{aff} = \text{True} \\ \text{NoLinearDealloc} \wedge \uparrow P & \text{if } \text{aff} = \text{False} \end{cases}$$

$$\text{holds_id}(id, \text{aff}) \triangleq \exists P \in \text{fProp}_0. \boxed{\circ [id \leftarrow \text{ag}(P)]}^{\text{Store}} * \text{holds}(P, \text{aff})$$

$$\text{assert}(i, \text{inv}^?) \triangleq \begin{cases} \text{Emp} & \text{if } \text{inv}^? = \perp \\ (\text{holds_id}(id, \text{aff}) * \boxed{\{i\}}^{\text{Dis}}) \vee \boxed{\{(ts, i)\}}^{\text{En}} & \text{if } \text{inv}^? = (id, \text{aff}, ts) \end{cases}$$

$$\text{wsat} \triangleq \left(\exists S. \boxed{\bullet S}^{\text{Store}} \right) * \left(\exists R. \boxed{\bullet R}^{\text{Reg}} * \bigstar_{i \in \text{dom}(R)} \text{assert}(i, R(i)) \right)$$

"holds" enforces that affine invariants only hold affine resources, and that linear invariants do not hold deallocation tokens for linear invariants

Concurrent Language

- User provides small-step operational semantics for expressions
- User provides definitions of *allowed* and *forbidden* stuckness
 - More customization of when code may be stuck

$$(\text{step} \text{ --- } \text{---}) \subseteq \text{ExecEnv} \times \text{Code} \times \text{State} \times \text{Code} \times \text{State} \times \text{Effect}$$
$$(\text{allowed_stuck} \text{ --- } \text{---}) \subseteq \text{ExecEnv} \times \text{Code} \times \text{State}$$
$$(\text{forbidden_stuck} \text{ --- } \text{---}) \subseteq \text{ExecEnv} \times \text{Code} \times \text{State}$$

Concurrent Language

- We define a concurrent language with threads, and asynchronous tasks in each thread
 - Asynchronous tasks follow "run-to-completion" scheduling (similar to JS)
 - In each thread, context switch occurs only at yield-points

$$\frac{\begin{array}{l} tp(t) = \langle a_{active}, ap \rangle \quad ap(a) = tk \\ (a = a_{active} \vee \text{can_yield}(ap(a_{active}))) \end{array} \quad \text{task_step } t \ p \ tp \ h \ tk \ h' \ tk' \ eff}{\text{thread_step } t \ p \ tp \ h \ \text{update_pool}(\text{handle_eff}(t, tp, eff), t, a, tk') \ h'}$$

Weakest Precondition

$$\text{wp}_{p;t;\mathcal{E}_1;\mathcal{E}_2} c \{v, \mathcal{E}. P\} \{c', \mathcal{E}'. Q\}$$

Program, Thread Identifier

Mask to be restored at yield-points, Current enabled mask

Code expression

Post-condition, Post-condition for coinduction^[1]

-
1. Lennard Gäher et al. Simuliris: A Separation Logic Framework for Verifying Concurrent Program Optimizations. POPL 2022 [↩](#)

Weakest Precondition

$wp_p \triangleq \mu wp_rec. \lambda t, \mathcal{E}_1, \mathcal{E}_2, c, \Phi, \Phi_{\text{coind}}. \forall tp, h.$

$state_inter(p, tp, h) \multimap_{\mathcal{E}_2} \emptyset ($

$\emptyset \Vdash_{\lambda_True} (\lambda \mathcal{E}_3. state_inter(p, tp, t) * \Phi_{\text{coind}}(c, \mathcal{E}_3)) \vee$ (COINDUCTIVE HYP. APP.)

$\emptyset \Vdash_{\circ(\mathcal{E}_1, c)} (\lambda \mathcal{E}_3. state_inter(p, tp, t) * \exists v. \ulcorner to_val(c) = v \urcorner \wedge \Phi(v, \mathcal{E}_3)) \vee$ (VALUE CASE)

$(\ulcorner red((p, tp, t), h, c) \vee allowed_stuck(p, tp, t) h c \urcorner \wedge$ (SAFETY CONDITION)

$\forall h', c', eff. \ulcorner step(p, tp, t) h c h' c' eff \urcorner \Rightarrow \emptyset \Vdash_{\circ(\mathcal{E}_1, c')} (\lambda \mathcal{E}_3.$ (STEP CASE)

$state_inter(p, handle_eff(t, tp, eff), h') * wp_rec(t, \mathcal{E}_1, \mathcal{E}_3, c', \Phi, \Phi_{\text{coind}})*$

$handle_new_task(...)))$

$\circ(\mathcal{E}_1, c) \triangleq \lambda \mathcal{E}_2. \begin{cases} \mathcal{E}_1 = \mathcal{E}_2 & \text{if } is_yield_point(c) \\ \{(ThShared, i) \mid (ThShared, i) \in \mathcal{E}_1\} \subseteq \mathcal{E}_2 & \text{otherwise} \end{cases}$

If c is a yield-point, we must enable all invariants (in \mathcal{E}_1).

Otherwise, we just need to enable all thread-shared invariants.

Object Calculus

$v \in \mathbf{Values} ::= \mathbf{null} \mid \mathbf{true} \mid \mathbf{false} \mid bv \mid \iota \mid a_t$ async task id

$e \in \mathbf{Exprs} ::=$

	v	value		x	variable
	$\mathbf{let} \ x = e; e'$	let binding		$e.f$	field access
	$e.f = e'$	field assign		$\mathbf{new} \ C$	new object
	$\mathbf{free} \ e$	free object		$e.m(e')$	method call
	$uop \ e$	unary operation		$e \ bop \ e'$	binary operation
	$\mathbf{if} \ e_1 \{e_2\} \ \mathbf{else} \ \{e_3\}$	if then else		$\mathbf{async} \ e$	create async task
	$\mathbf{fork} \ e$	create thread		$\mathbf{await} \ e$	await task
	$\mathbf{lock} \ e$	lock object		$\mathbf{unlock} \ e$	unlock object

Both examples verified

```
let queue = new Queue.init(); queue.use();
```

```
let consumer =  
  async { new Consumer.init(queue) };  
let producer =  
  async { new Producer.init(queue) };
```

or

```
let consumer =  
  fork { new Consumer.init(queue) };  
let producer =  
  fork { new Producer.init(queue) };
```

```
await producer; await consumer; queue.isEmpty()
```

$$\{\text{Emp}\} \text{main } \{v, \mathcal{E}. \ulcorner v = \text{true} \wedge \mathcal{E} = \text{imask}(0) \urcorner \wedge \text{Emp}\}_{p;0;\text{imask}(0);\text{imask}(0)}$$

Proof strategy

1. Use counting permissions instead of fractional ones
2. Allocate invariant that stores useCount opening tokens for itself
3. Give the full deallocation token to the consumer
4. Give one opening token to the producer

$$\text{QueueInv}(i, \iota) \triangleq \exists uc \ l. \text{Queue}(1, \iota, uc, l) * \text{OPermToken}(i, uc)$$

$\{\text{QueueDPerm}(i, t, 0, \iota)\}$ **new** Consumer.init(ι) $\{-, \mathcal{E}. \ulcorner \mathcal{E} = \text{imask}(t) \urcorner \wedge \text{Queue}(1, \iota, 0, [])\}_{p;t;\text{imask}(t);\text{imask}(t)}$

$\{\text{QueueOPerm}(i, t, -1, \iota)\}$ **new** Producer.init(ι) $\{-, \mathcal{E}. \ulcorner \mathcal{E} = \text{imask}(t) \urcorner \wedge \text{Emp}\}_{p;t;\text{imask}(t);\text{imask}(t)}$

Results

```
Theorem main_safe tp' h' :  
  tpool_steps prog {[ $\theta := (\theta, \{[\theta := \text{Start main } ]\})$ ]}  $\emptyset$  tp' h' ->  
  ~ forbidden_stuck_tp prog tp' h'.
```

Proof. ... **Qed.**

```
Theorem main_sound tp' h' v :  
  tpool_steps prog {[ $\theta := (\theta, \{[\theta := \text{Start main } ]\})$ ]}  $\emptyset$  tp' h' ->  
  get_finished_task tp'  $\theta$   $\theta$  = Some v ->  
  v = v_True.
```

Proof. ... **Qed.**

```
Theorem main_leak_free tp' h' :  
  tpool_steps prog {[ $\theta := (\theta, \{[\theta := \text{Start main } ]\})$ ]}  $\emptyset$  tp' h' ->  
  terminated_tp tp' ->  
  map_Forall ( $\lambda$  _ obj, let (C, F, lk) := obj in get_aff prog C = true) h'.
```

Proof. ... **Qed.**

No forbidden configurations are reached

If the program terminates, it yields `true` and all objects in the heap are affine

Thank you!

jd.mota@campus.fct.unl.pt

jdmota.github.io

- Affine and linear resources/invariants
- Thread-local/shared invariants
- Weakest-precondition for concurrent language ensuring safety and leak freedom