François Pottin Jean-Marie Madiot



A Separation Logic for Heaf Space under GC

Reasoning about Heap Space

We wish to verify a program's *heap space* usage,

- using *separation logic*,
- by viewing heap space as a *resource*...



Idea 1. Following Hofmann (1999), let \diamond 1 represent one *space credit*. Allocation consumes credits; deallocation produces credits.

$$\left\{ \diamondsuit{} \text{size}(b) \right\} \quad x := \text{alloc}(b) \left\{ x \mapsto b \right\}$$
$$\left\{ x \mapsto b \right\} \quad \text{free}(x) \quad \left\{ \diamondsuit{} \text{size}(b) \right\}$$

A function's space requirement is visible in its specification.



In the presence of GC, what Happens?

In the presence of GC,

- deallocation becomes *implicit*,
- so we lose the ability to recover space credits while reasoning.



Idea 2. Switch to a *logical deallocation* operation:

A ghost update \Rightarrow *consumes* an assertion and *produces* an assertion. This marries

- manual reasoning about memory at verification time
- automatic management of memory at runtime.





Is logical deallocation sound?

$$x \mapsto b \equiv$$
 (b)

It does have a few good properties: *no double-free, no use-after-free.* Because $x \mapsto b$ is consumed,

- a block cannot be logically deallocated twice;
- a block cannot be accessed after it has been logically deallocated.



Unfortunately, logical deallocation in this form is not sound.



Introducing logical deallocation creates a distinction between

- the *logical heap* that the programmer keeps in mind,
- the *physical heap* that exists at runtime.



The following situation is problematic.

The programmer has logically deallocated a block and obtained \diamond 3,



but this block is *reachable* and cannot be reclaimed by the GC. We have 3 space credits but *no free space* in the physical heap!



To avoid this problem, we must *restrict logical deallocation*:

• A reachable block must not be deallocated.

In the contrapositive,

- A block should be logically deallocatable only if it is unreachable,
- so the GC can reclaim this block,
- so the logical and physical heaps remain synchronized.



The Desired Global Invariant

The logical and physical heaps coincide on their reachable fragments.





How do we restrict logical deallocation?

- We want to disallow deallocating a *reachable* block,
- but Separation Logic lets us reason about *ownership*.
- Reachability is a *nonlocal* property.



Idea 3. Following Kassios and Kritikos (2013),

- we keep track of the predecessors of every block.
- If a block has no predecessor, *then* it is unreachable,
- therefore it can be logically deallocated.



Points-To and Pointed-By Assertions

In addition to *points-to*, we use *pointed-by* assertions:





Logical Deallocation

We get a sound logical deallocation axiom:

$$x \mapsto b * x \leftarrow \phi \Rightarrow Aize(b)$$

This axiom deallocates one block.

There is also a *bulk logical deallocation* axiom.



We want the pointers from the stack(s) to the heap to be explicit,

- so the operational semantics views them as GC roots,
- so our predecessor-tracking logic keeps track of them.

Idea 4. Use a low-level calculus where stack cells are explicit.



Roadmap

1 A Glimpse of SpaceLang

2 A Glimpse of the Reasoning Rules

Specification of a Stack





SpaceLang is imperative. An *instruction i* does not return a value.

skip	по-ор	$*\varrho = \text{alloc } n$	heap allocation
i; i	sequencing	$*\varrho = [*\varrho + o]$	heap load
if $*\varrho$ then <i>i</i> else <i>i</i>	conditional	$[*\varrho + o] = *\varrho$	heap store
$*\varrho(\vec{\varrho})$	procedure call	$*\varrho = (*\varrho = = *\varrho)$	address comparison
$*\varrho = v$	constant load	alloca x in <i>i</i>	stack allocation
$*\varrho = *\varrho$	move	fork $*\varrho$ as x in i	thread creation

The operands of every instruction are stack cells ρ .

There is *no heap deallocation* instruction.



A small-step operational semantics, with a few unique features:

- *Garbage collection* takes place before every reduction step.
- The GC *roots* are the stack cells.
- Heap allocation *fails* if the heap size exceeds a fixed limit *S*.



Roadmap

1 A Glimpse of SpaceLang

2 A Glimpse of the Reasoning Rules

Specification of a Stack

4 Conclusion



Heap allocation *consumes space credits*.

ALLOC
diamondi
continued
$$\left\{ \begin{array}{c} \diamond size(()^n) \\ s \mapsto \langle v \rangle \\ v \leftarrow q L \end{array} \right\} \ *s = alloc n$$

$$\left\{ \begin{array}{c} \ell \mapsto ()^n \\ \ell \leftarrow \{s\} \\ \exists \ell \\ s \mapsto \langle \ell \rangle \\ v \leftarrow q L \setminus \{s\} \\ \end{bmatrix} \right\}$$
updated stack cell edge deletion

Points-to and pointed-by assertions for the new location appear.

One pointer to the value v is *deleted*.



Heap Store

Reasoning about a heap store involves some administration...



One pointer to v is *deleted*; one pointer to v' is *created*.



Logical Deallocation

Logical deallocation of a block is a *ghost operation:*





Theorem (Soundness)

If $\{\diamond S\}$ i $\{True\}$ holds, then, executing i in an empty store cannot lead to a situation where a thread is stuck.

If, under a precondition of S space credits, the code can be verified, *then* its *live* heap space cannot exceed S.

This holds *regardless of the value of S* (the heap size limit). Furthermore, the reasoning rules are *independent* of S.

The rules allow *compositional reasoning* about space.



Roadmap

1 A Glimpse of SpaceLang

2 A Glimpse of the Reasoning Rules

3 Specification of a Stack

4 Conclusion



The user may define *custom* (simplified) predecessor tracking disciplines. For example, sometimes, *counting* predecessors is enough.

$$v \leftrightarrow n \triangleq \exists L. (v \leftrightarrow_1 L \star |L| = n)$$

Edge addition and deletion increment and decrement n.



Creating a stack consumes 4 space credits.

$$\begin{cases} f \mapsto \langle create \rangle \\ stack \mapsto \langle () \rangle \\ \diamond 4 \end{cases} * f(stack) \begin{cases} f \mapsto \langle create \rangle \\ \exists \ell. \quad stack \mapsto \langle \ell \rangle \\ isStack \ \ell \ [] \star \ell \leftarrow 1 \end{cases}$$

We get unique ownership of the stack and *we have the sole pointer* to it.



Pushing

Pushing consumes 4 space credits.

$$\begin{cases} f \mapsto \langle push \rangle \\ stack \mapsto \langle \ell \rangle \\ elem \mapsto \langle v \rangle \\ \diamond 4 \star isStack \ \ell \ vs \\ v \leftarrow n \end{cases} \ast f(stack, elem) \begin{cases} f \mapsto \langle push \rangle \\ stack \mapsto \langle \ell \rangle \\ elem \mapsto \langle v \rangle \\ isStack \ \ell \ (v :: vs) \\ v \leftarrow n + 1 \end{cases}$$

The value v receives one more antecedent.



Popping

Popping frees up 4 space credits.

$$\begin{pmatrix} f \mapsto \langle pop \rangle \\ stack \mapsto \langle \ell \rangle \\ elem \mapsto \langle () \rangle \\ isStack \ \ell \ (v :: vs) \\ v \leftarrow n \end{pmatrix} * f(stack, elem) \begin{cases} f \mapsto \langle pop \rangle \\ stack \mapsto \langle \ell \rangle \\ elem \mapsto \langle v \rangle \\ \diamond 4 \star isStack \ \ell \ vs \\ v \leftarrow n \end{cases}$$

The number of predecessors of v is unchanged, because the out-parameter *elem* receives a pointer to it.



Logically deallocation of the stack, a *ghost operation*, is part of the API. It requires proving that the stack has *zero predecessors*.

$$\begin{cases} isStack \ \ell \ vs \ \star \ \ell \ \leftarrow \ 0 \\ & \\ & \\ (v,n) \in vns \end{cases} \Rightarrow I \begin{cases} \diamond (4 + 4 \times |vs|) \\ & \\ & \\ (v,n) \in vns \end{cases} \end{cases}$$

It frees up a linear number of space credits.



Roadmap

1 A Glimpse of SpaceLang

2 A Glimpse of the Reasoning Rules

Specification of a Stack





Summary of Contributions

A sound logic to reason about heap space usage in the presence of GC. Our main insights:

- Allocation consumes *space credits* $\diamond n$.
- Logical deallocation, a ghost operation, produces space credits.
- Logical dellocation requires predecessor tracking, which we perform via pointed-by assertions v ← L.



Currently, predecessor tracking requires *heavy bookkeeping*. We are investigating

- a more flexible *deferred* logical edge deletion mechanism;
- coarse-grained predecessor tracking based on *islands*;
- simpler / more automated tracking of roots;
- reasoning directly about call-by-value λ -calculus.





5 Syntax, Semantics of SpaceLang

⑥ Reasoning Rules of SL◊

O Specification of List Copy



Values, Blocks, Stores

Memory locations: $\ell, c, r, s \in \mathcal{L}$.

Values include constants, memory locations, and *closed procedures:*

$$v ::= () \mid k \mid \ell \mid \lambda \vec{x}.i$$

Memory blocks include *heap tuples*, *stack cells*, and deallocated blocks:

$$b ::= ec{v} \mid \langle v
angle \mid \mathbf{4}$$

A *store* maps locations to blocks, encompassing the heap and stack(s). The *size* of a block:

$$size(\vec{v}) = 1 + |\vec{v}|$$
 $size(\langle v \rangle) = size(f) = 0$

The size of the store is the sum of the sizes of all blocks.



A *reference* is a variable or a (stack) location and denotes a *stack cell*.

$$\varrho ::= x \mid c$$

SpaceLang uses *call-by-reference*.

A variable denotes a closed reference, *not* a closed value as is usual. The operational semantics involves substitutions [c/x]. This preserves the property that *the code never points to the heap*.

The roots of the garbage collection process are the stack cells.



SpaceLang is imperative. An *instruction i* does not return a value.

skip	no-op	$*\varrho = \text{alloc } n$	heap allocation
i; i	sequencing	$*\varrho = [*\varrho + o]$	heap load
if $*\varrho$ then <i>i</i> else <i>i</i>	conditional	$[*\varrho + o] = *\varrho$	heap store
$*\varrho(\vec{\varrho})$	procedure call	$*\varrho = (*\varrho == *\varrho)$	address comparison
$*\varrho = v$	constant load	alloca x in <i>i</i>	stack allocation
$*\varrho = *\varrho$	move	alloca <i>c</i> in <i>i</i>	active stack cell
		fork $*\varrho$ as x in i	thread creation

The operands of every instruction are stack cells (ρ).

There is no deallocation instruction for heap blocks.



Operational Semantics: Heap Allocation

We fix a *maximum heap size S*.

Heap allocation fails if the heap size exceeds S.

STEPALLOC

$$\sigma' = [\ell += ()^{n}]\sigma$$

$$\frac{size(\sigma') \le S}{s = alloc n / \sigma \longrightarrow skip / \sigma''}$$

S is a parameter of the operational semantics,

but the reasoning rules of SL \diamond are independent of S.



Operational Semantics: Stack Allocation

The dynamic semantics of stack allocation is in *three steps*:

STEP ALLOCA ENTRY $\sigma' = [c += \langle () \rangle] \sigma$ alloca x in $i / \sigma \longrightarrow$ alloca c in $[c/x]i / \sigma'$ alloca c in skip $/ \sigma \longrightarrow$ skip $/ \sigma'$

STEPALLOCAEXIT $\sigma(c) = \langle v \rangle \qquad \sigma' = [c := \mathbf{i}]\sigma$

Evaluation contexts: K ::= [] | K; i | alloca c in K.



To complete the definition of the operational semantics,

- allow *garbage collection* before every reduction step.
 - $\sigma \boxdot \sigma'$ holds if
 - the stores σ and σ' have the same domain;
 - for every ℓ in this domain, either $\sigma'(\ell) = \sigma(\ell)$, or ℓ is unreachable in σ and $\sigma'(\ell) = \mathbf{1}$.
- allow thread interleavings (comes for free with Iris).



Complete Operational Semantics

StepSeqSkip	STEPIF $\sigma(r) = \langle k \rangle$	ŝ	$\begin{array}{l} \text{StepCall} \\ \sigma(\mathbf{r}) = \langle \lambda \vec{x}. i \rangle & \vec{x} = \vec{s} \end{array}$	
skip; $i / \sigma \longrightarrow i / \sigma$	$if * r then i_1 else i_2 / \sigma \longrightarrow k \neq 0$	$? i_1 : i_2 / \sigma$	$\frac{(\vec{s}) \ / \ \sigma \longrightarrow [\vec{s}/\vec{x}]i \ / \ \sigma}{\ast r(\vec{s}) \ / \ \sigma \longrightarrow [\vec{s}/\vec{x}]i \ / \ \sigma}$	
$STEPCONST \sigma' = \langle s := v \rangle \sigma pointers(v) = \emptyset \overline{*s = v / \sigma \longrightarrow \text{skip} / \sigma'}$	$\frac{\text{STEPMOVE}}{\sigma(r) = \langle v \rangle} \\ \frac{\sigma' = \langle s := v \rangle \sigma}{*s = *r / \sigma \longrightarrow \text{skip} / \sigma}$	$\frac{\text{STEPALLO}}{\sigma'} \qquad \frac{\text{size}(\sigma')}{*s = \text{al}}$	$\begin{aligned} \sigma' &= [\ell += ()^n] \sigma \\ &\leq S \sigma'' = \langle s := \ell \rangle \sigma' \\ &\text{loc } n \ / \ \sigma \longrightarrow \text{skip} \ / \ \sigma'' \end{aligned}$	
$\frac{\sum_{\sigma(r) = \langle \ell \rangle}^{\text{STEPLOAD}} \sigma(\ell)}{\vec{v}(o) = v} \frac{\sigma(\ell)}{*s = [*r + o]}$	$\sigma' = \vec{v} 0 \le o < \vec{v} $ $\sigma' = \langle s := v \rangle \sigma$ $\sigma \longrightarrow \text{skip} / \sigma'$	$\frac{\substack{\text{STEPSTORE}\\\sigma(r) = \langle v \rangle \sigma(s) = \\ 0 \le o < \vec{v} \sigma' = \\ [*s + o] = *r / \sigma}$	$ \begin{array}{c} = \langle \ell \rangle & \sigma(\ell) = \vec{v} \\ = [\ell := [o := v] \vec{v}] \sigma \\ \hline \sigma \longrightarrow \text{skip} / \sigma' \end{array} $	
STEPLOCEQ $\sigma(r_1) = \langle \ell_1 \rangle$ $\sigma' = \langle s := (\ell_1)$ $s = (*r_1 = *r_2)$	$\sigma(r_2) = \langle \ell_2 \rangle \qquad S$ = ℓ_2 ? 1 : 0) $\rangle \sigma$ / $\sigma \longrightarrow \text{skip} / \sigma'$	$\frac{\sigma' = [c + =]}{\log a \times \ln i / \sigma} \xrightarrow{\sigma' = [c + =]}{\operatorname{all}}$	$\langle () \rangle] \sigma$ bca c in [c/x]i / σ'	
$\frac{STEPALLOCAEXIT}{\sigma(c) = \langle v \rangle \sigma' = [c := \P]}$ alloca c in skip / $\sigma \longrightarrow$ skip /	$\frac{\sigma}{\sigma'} \qquad \frac{\frac{S \text{TEPFORK}}{\sigma(r) = \langle v \rangle} \sigma' = [r]}{\frac{\sigma(r) = \langle v \rangle}{\text{fork } * r \text{ as } x \text{ in } i / \sigma}}$	$\frac{(i)[c + = \langle v \rangle]\sigma}{\longrightarrow \text{skip } / \sigma'}$ $\frac{\sigma'}{ \text{loca } c \text{ in } [c/x]i}$	$\frac{\begin{array}{c} \text{STEPCONTEXT} \\ i \ / \ \sigma \longrightarrow i' \ / \ \sigma' \\ \hline \\ \hline \\ \hline \\ \hline \hline \\ \hline \hline \\ \hline \hline \\ \hline \\ \hline $	

Roadmap

5 Syntax, Semantics of SpaceLang

6 Reasoning Rules of SL◊

O Specification of List Copy



Heap allocation *consumes space credits*.

ALLOC
diamonds
consumed
$$\binom{\diamond size(()^n)}{s \mapsto \langle v \rangle}$$

 $v \leftarrow q L$ $\ast s = alloc n$ $\begin{cases} \ell \mapsto ()^n \\ \ell \leftarrow \{s\} \\ \vdots \\ v \leftarrow q L \setminus \{s\} \\ \end{cases}$
updated stack cell edge deletion

Points-to and pointed-by assertions for the new location appear. One pointer to the value v is *deleted*. (This aspect is optional.)



Heap Store

Writing a heap cell is simple... but involves some administration.



One pointer to v is deleted; one pointer to v' is *created*.



Stack Allocation

A points-to assertion for the new stack cell exists throughout its lifetime.



No pointed-by assertion is provided. (A design choice.)

• No pointers (from the heap or stack) to the stack.



Logical Deallocation

Logical deallocation of a block is a *ghost operation*:





Deletion of deallocated predecessors can be *deferred*:

A key rule: if L' is empty, then v becomes eligible for deallocation.



Bulk Logical Deallocation

A group that is *closed under predecessors* can be deallocated at once:



The rules for constructing a "cloud" (omitted) are straightforward.



Points-to and pointed-by assertions can be *split* and *joined*.

$$l \mapsto_{q_{1}+q_{2}} b \equiv l \mapsto_{q_{1}} b * l \mapsto_{q_{2}} b$$

$$v \leftarrow_{q_{1}+q_{2}} L_{n} \forall L_{2} \equiv v \leftarrow_{q_{1}} L_{n} * v \leftarrow_{q_{2}} L_{2}$$

$$v \leftarrow_{q} L \longrightarrow v \leftarrow_{q} L' \qquad i \xi L \subseteq L'$$

$$l \mapsto_{q} b * l' \leftarrow_{1} L \Rightarrow_{1} l \mapsto_{q} b * l' \leftarrow_{1} L *$$

$$[l' & pointers(b) \leq l \leq L]$$

Pointed-by assertions are *covariant*.

Points-to and pointed-by assertions can be *confronted*.



Space credits can be *split* and *joined*.

True
$$\Longrightarrow \diamond 0$$

 $\diamond (m_1 + m_2) \implies \diamond m_1 * \diamond m_2$



Roadmap

5 Syntax, Semantics of SpaceLang

⑥ Reasoning Rules of SL◊

7 Specification of List Copy



Each cell owns the next cell and possesses the sole pointer to it.

$$\begin{array}{ccc} \mathsf{isList}\ \ell\ [] &\triangleq & \ell \mapsto [0]\\ \mathsf{isList}\ \ell\ (\mathsf{v}::\mathsf{vs}\) &\triangleq & \exists \ell'.\ \ell \mapsto [1;\ \mathsf{v};\ \ell'] \star \ell' \leftrightarrow 1 \star \mathsf{isList}\ \ell'\ \mathsf{vs} \end{array}$$

Let's now have a look at *list copy* and its spec. (Fasten seatbelts!)



List Copy in SpaceLang

$$copy \triangleq \lambda(self, dst, src).$$

alloca tag in $*tag = [*src + 0];$
if $*tag$ then
alloca $head$ in $*head = [*src + 1];$
alloca $tail$ in $*tail = [*src + 2];$
 $*src = ();$
alloca dst' in $*self(self, dst', tail);$
 $*dst = alloc 3;$
 $[*dst + 0] = *tag;$
 $[*dst + 1] = *head;$
 $[*dst + 2] = *dst'$
else
 $*src = ();$
 $*dst = alloc 1;$
 $[*dst + 0] = *tag$

- read the list's tag
- if this is a cons cell, then
- read the list's head
- read the list's tail
- clobber this root
- copy the list's tail
- allocate a new cons cell
- and initialize it
- this must be a nil cell
- clobber this root
- allocate a new nil cell
- and initialize it



Specification of List Copy

The case m = 1, where we have *the sole pointer* to the list, is special.

$$\begin{cases} f \mapsto \langle copy \rangle \star dst \mapsto \langle () \rangle \star src \mapsto \langle \ell \rangle \\ isList \ \ell \ vs \ \star \ \ell \leftarrow m \\ m = 1 ? \diamond 0 : \diamond (2 + 4 \times |vs|) \\ \forall v \in vs. \ \exists n. \ (v, n) \in vns \\ \star_{(v,n) \in vns} v \leftarrow n \end{cases}$$
 meed no space or linear space

 $\exists \ell'. \begin{cases} f \mapsto \langle copy \rangle * dst \mapsto \langle \ell' \rangle * src \mapsto \langle () \rangle \\ m = 1 ? True : (isList \ell vs * \ell \leftrightarrow m - 1) \\ isList \ell' vs * \ell' \leftrightarrow 1 \\ *_{(v,n) \in vns} v \leftrightarrow n + (m = 1 ? 0 : v $ vs) \end{cases} \xrightarrow{\text{ous. Bit is deallocated} \\ \underbrace{\mathfrak{Q}}_{\text{reserved}} \xrightarrow{\text{preserved}} \\ \underbrace{\mathfrak{Q}}_{\text{reserved}} \xrightarrow{\text{reserved}} \\ \underbrace{\mathfrak{Q}}_{\text{reserved}} \xrightarrow{\text{reserved}} \\ \underbrace{\mathfrak{Q}}_{\text{reserved}} \xrightarrow{\text{reserved}} \xrightarrow{\text{reserved}} \\ \underbrace{\mathfrak{Q}}_{\text{reserved}} \xrightarrow{\text{reserved}} \xrightarrow{\text{reserved}} \\ \underbrace{\mathfrak{Q}}_{\text{reserved}} \xrightarrow{\text{reserved}} \xrightarrow{\text{rese$