Modular Metatheory for Abstract Interpreters

Abstract

The design and implementation of static analyzers have becoming increasingly systematic. In fact, design and implementation have remained seemingly on the verge of full mechanization for several years. A stumbling block in full mechanization has been the ad hoc nature of soundness proofs accompanying each analyzer. While design and implementation is largely systematic, soundness proofs can change significantly with (apparently) minor changes to the semantics and analyzers themselves. We finally reconcile the systematic construction of static analyzers with their proofs of soundness via a mechanistic Galois-connection-based metatheory for static analyzers.

1. Introduction

Writing abstract interpreters is hard. Writing proofs about abstract interpreters is extra hard. Modern practice in whole-program analysis requires multiple iterations in the design space of possible analyses. As we explore the design space of abstract interpreters, it would be nice if we didn't need to reprove all the properties we care about. What we lack is a reusable meta-theory for exploring the design space of *correct-by-construction* abstract interpreters.

We propose a compositional meta-theory framework for general purpose static analysis. Our framework gives the analysis designer building blocks for building correct-by-construction abstract interpreters. These building blocks are compositional, and they carry both computational and correctness properties of an analysis. For example, we are able to tune the flow and path sensitivities of an analysis in our framework with no extra proof burden. We do this by capturing the essential properties of flow and path sensitivities into plug-and-play components. Comparably, we show how to design an analysis to be correct for all possible instantiations to flow and path sensitivity.

To achieve compositionality, our framework leverages monad transformers as the fundamental building blocks for an abstract interpreter. Monad transformers snap together to form a single monad which drives interpreter execution. Each piece of the monad transformer stack corresponds to either an element of the semantics' state space or a nondeterminism effect. Variations in the transformer stack to give rise to different path and flow sensitivities for the analysis. Interpreters written in our framework are proven correct w.r.t. all possible monads, and therefore to each choice of path and flow sensitivity.

The monad abstraction provides the computational and proof properties for our interpreters, from the monad operators and laws respectively. Monad transformers are monad composition function; they consume and produce monads. We strengthen the monad transformer interface to require that the resulting monad have a relationship to a state machine transition space. We prove that a small set of monads transformers that meet this stronger interface can be used to write monadic abstract interpreters.

1.1 Contributions:

Our contributions are:

- A compositional meta-theory framework for building correct-by-construction abstract interpreters. This framework is built using a restricted class of monad transformers.
- An isolated understanding of flow and path sensitivity for static analysis. We understand this spectrum as mere variations in the order of monad transformer composition in our framework.

1.2 Outline

We will demonstrate our framework by example, walking the reader through the design and implementation of an abstract interpreter. Section X gives the concrete semantics for a small functional language. Section X shows the full definition of a highly parameterized monadic interpreter. Section [X][Recovering Concrete and Abstract Interpreters] shows how to recover concrete and abstract interpreters. Section X shows how to manipulate the path and flow sensitivity of the interpreter through varyations in the monad. Section X demonstrates our compositional meta-theory framework built on monad transformers.

2. Semantics

To demonsrate our framework we design an abstract interpreter for a simple applied lambda calculus: IF.

$$
\begin{array}{rll}
i & \in \mathbb{Z} & \\
x & \in \mathit{Var} & \\
a & \in \mathit{Atom} & ::= i \mid x \mid \underline{\boldsymbol{\lambda}}(x).e \\
\oplus & \in \mathit{IOp} & ::= + \mid - \\
\odot & \in \mathit{Op} & ::= \oplus \mid @ \\
e & \in \mathit{Exp} & ::= a \mid e \odot e \mid \underline{\textbf{if0}}(e)\{e\}\{e\}
\end{array}
$$

IF extends traditional lambda calculus with integers, addition, subtration and conditionals. We use the operator @ as explicit syntax for function application. This allows for *Op* to be a single syntactic class for all operators and simplifies the presentation.

Before designing an abstract interpreter we first specify a formal semantics for IF. Our semantics makes allocation explicit and separates values and continuations into separate stores. Our approach to analysis will be to design a configurable interpreter that is capable of mirroring these semantics.

The state space Σ for IF is a standard CESK machine augmented with a separate store for continuation values:

$$
\begin{array}{rll}
\tau & \in \mathit{Time} & ::= \mathbb{Z} \\
l & \in \mathit{Addr} & ::= \mathit{Var} \times \mathit{Time} \\
\rho & \in \mathit{Env} & ::= \mathit{Var} \rightharpoonup \mathit{Addr} \\
\sigma & \in \mathit{Store} & ::= \mathit{Addr} \rightharpoonup \mathit{Val} \\
c & \in \mathit{Clo} & ::= \langle \underline{\boldsymbol{\lambda}}(x).e, \rho \rangle \\
v & \in \mathit{Val} & ::= i \mid c \\
\kappa l & \in \mathit{KAddr} & ::= \mathit{Time} \\
\kappa\sigma & \in \mathit{KStore} & ::= \mathit{KAddr} \rightharpoonup \mathit{Frame} \times \mathit{KAddr} \\
fr & \in \mathit{Frame} & ::= \langle \square \odot e \rangle \mid \langle v \odot \square \rangle \mid \langle \underline{\textbf{if0}}(\square)\{e\}\{e\} \rangle \\
\varsigma & \in \Sigma & ::= \mathit{Exp} \times \mathit{Env} \times \mathit{Store} \times \mathit{KAddr} \times \mathit{KStore}
\end{array}
$$

The semantics of atomic terms is given denotationally with the denotation function $A[\![\_, \_, \_]\!]$:

$$
\begin{array}{rl}
A[\![\_, \_, \_]\!] & \in \mathit{Env} \times \mathit{Store} \times \mathit{Atom} \rightharpoonup \mathit{Val} \\
A[\![\rho, \sigma, i]\!] & := i \\
A[\![\rho, \sigma, x]\!] & := \sigma(\rho(x)) \\
A[\![\rho, \sigma, \underline{\boldsymbol{\lambda}}(x).e]\!] & := \langle \underline{\boldsymbol{\lambda}}(x).e, \rho \rangle \\
\delta[\![\_, \_, \_]\!] & \in \mathit{IOp} \times \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z} \\
\delta[\![+, i_1, i_2]\!] & := i_1 + i_2 \\
\delta[\![-, i_1, i_2]\!] & := i_1 - i_2
\end{array}
$$

The semantics of compound expressions are given relationally via the step relation $\_ \rightsquigarrow \_$:

$$
\begin{array}{l}
\_ \rightsquigarrow \_ \ \in \ \mathcal{P}(\Sigma \times \Sigma) \\
\langle e_1 \odot e_2, \rho, \sigma, \kappa l, \kappa\sigma, \tau \rangle \rightsquigarrow \langle e_1, \rho, \sigma, \tau, \kappa\sigma', \tau + 1 \rangle \\
\quad \text{where} \ \ \kappa\sigma' := \kappa\sigma[\tau \mapsto \langle \square \odot e_2 \rangle :: \kappa l] \\
\langle a, \rho, \sigma, \kappa l, \kappa\sigma, \tau \rangle \rightsquigarrow \langle e, \rho, \sigma, \tau, \kappa\sigma', \mathit{tick}(\tau) \rangle \\
\quad \text{where} \\
\qquad \langle \square \odot e \rangle :: \kappa l' := \kappa\sigma(\kappa l) \\
\qquad \kappa\sigma' := \kappa\sigma[\tau \mapsto \langle A[\![\rho, \sigma, a]\!] \odot \square \rangle :: \kappa l'] \\
\langle a, \rho, \sigma, \kappa l, \kappa\sigma, \tau \rangle \rightsquigarrow \langle e, \rho'', \sigma', \kappa l', \kappa\sigma, \tau + 1 \rangle \\
\quad \text{where} \\
\qquad \langle \langle \underline{\boldsymbol{\lambda}}(x).e, \rho' \rangle \ @ \ \square \rangle :: \kappa l' := \kappa\sigma(\kappa l) \\
\qquad \sigma' := \sigma[(x, \tau) \mapsto A[\![\rho, \sigma, a]\!]] \\
\qquad \rho'' := \rho'[x \mapsto (x, \tau)] \\
\langle i_2, \rho, \sigma, \kappa l, \kappa\sigma, \tau \rangle \rightsquigarrow \langle i, \rho, \sigma, \kappa l', \kappa\sigma, \tau + 1 \rangle \\
\quad \text{where} \\
\qquad \langle i_1 \oplus \square \rangle :: \kappa l' := \kappa\sigma(\kappa l) \\
\qquad i := \delta[\![\oplus, i_1, i_2]\!] \\
\langle i, \rho, \sigma, \kappa l, \kappa\sigma, \tau \rangle \rightsquigarrow \langle e, \rho, \sigma, \kappa l', \kappa\sigma, \tau + 1 \rangle \\
\quad \text{where} \\
\qquad \langle \underline{\textbf{if0}}(\square)\{e_1\}\{e_2\} \rangle :: \kappa l' := \kappa\sigma(\kappa l) \\
\qquad e := e_1 \ \ \text{when} \ \ i = 0 \\
\qquad e := e_2 \ \ \text{when} \ \ i \neq 0
\end{array}
$$

Our abstract intepreter will support abstract garbage collection [CITE], the concrete analogue of which is just standard garbage collection. Garbage collection is defined with a reachability function *R* which computes the transitively reachable address from (ρ, e) in σ:

$$
\begin{array}{l}
R[\_] \ \in \ \mathit{Store} \to \mathit{Env} \times \mathit{Exp} \to \mathcal{P}(\mathit{Addr}) \\
R[\sigma](\rho, e) := \mu(X). \\
\quad R_0(\rho, e) \cup X \cup \{l' \mid l' \ \in \ \textit{R-Val}(\sigma(l)) \ ; \ l \ \in \ X\}
\end{array}
$$

We write $\mu(X).f(X)$ as the least-fixed-point of a function f. This definition uses two helper functions: R_0 for computing the initial reachable set adn *R-Val* for computing addresses reachable from addresses.

$$
\begin{array}{l}
R_0 \ \in \ \mathit{Env} \times \mathit{Exp} \to \mathcal{P}(\mathit{Addr}) \\
R_0(\rho, e) := \{\rho(x) \mid x \ \in \ \mathit{FV}(e)\} \\
\textit{R-Val} \ \in \ \mathit{Val} \to \mathcal{P}(\mathit{Addr}) \\
\textit{R-Val}(i) := \{\} \\
\textit{R-Val}(\langle \underline{\boldsymbol{\lambda}}(x).e, \rho \rangle) := \{\rho(x) \mid y \ \in \ \mathit{FV}(\underline{\boldsymbol{\lambda}}(x).e)\}
\end{array}
$$

FV is the standard recursive definition for computing free variables of an expression:

$$
\begin{aligned}
&\mathit{FV} \in \mathit{Exp} \to \mathcal{P}(\mathit{Var}) \\
&\mathit{FV}(x) \coloneqq \{x\} \\
&\mathit{FV}(i) \coloneqq \{\} \\
&\mathit{FV}(\underline{\boldsymbol{\lambda}}(x).e) \coloneqq \mathit{FV}(e) - \{x\} \\
&\mathit{FV}(e_1 \odot e_2) \coloneqq \mathit{FV}(e_1) \cup \mathit{FV}(e_2) \\
&\mathit{FV}(\underline{\mathbf{if0}}(e_1)\{e_2\}\{e_3\}) \coloneqq \mathit{FV}(e_1) \cup \mathit{FV}(e_2) \cup \mathit{FV}(e_3)
\end{aligned}
$$

Analagously, KR is the set of transitively reachabel continuation addresses in $\kappa\sigma$:

$$
\begin{aligned}
&KR[\_] \in \mathit{KStore} \to \mathit{KAddr} \to \mathcal{P}(\mathit{KAddr}) \\
&KR[\kappa\sigma](\kappa l) \coloneqq \mu(k\theta).\kappa\theta_0 \cup \kappa\theta \cup \{\pi_2(\kappa\sigma(\kappa l)) \mid \kappa l \in \kappa\theta\}
\end{aligned}
$$

Our final semantics is given via the step relation $\_ \rightsquigarrow^{gc} \_$ which nondeterministically either takes a semantic step or performs garbage collection.

$$
\begin{aligned}
&\_ \rightsquigarrow^{gc} \_ \in \mathcal{P}(\Sigma \times \Sigma) \\
&\varsigma \rightsquigarrow^{gc} \varsigma' \\
&\quad \text{where } \varsigma \rightsquigarrow \varsigma' \\
&\langle e, \rho, \sigma, \kappa l, \kappa\sigma, \tau\rangle \rightsquigarrow^{gc} \langle e, \rho, \sigma', \kappa l, \kappa\sigma, \tau\rangle \\
&\quad \text{where} \\
&\qquad \sigma' \coloneqq \{l \mapsto \sigma(l) \mid l \in R[\sigma](\rho, e)\} \\
&\qquad \kappa\sigma' \coloneqq \{\kappa l \mapsto \kappa\sigma(\kappa l) \mid \kappa l \in KR[\kappa\sigma](\kappa l)\}
\end{aligned}
$$

An execution of the semantics is states as the least-fixed-point of a collecting semantics:

$$\mu(X).\{\varsigma_0\} \cup X \cup \{\varsigma' \mid \varsigma \rightsquigarrow^{gc} \varsigma' \,;\, \varsigma \in X\}$$

We will justify our analyses as sound approximations of this collecting semantics.

3. Monadic Interpreter

In this section we design a monadic interpreter for the IF language which is also parameterizedin AAM[CITE] style. When finished, we will be able to recover a concrete interpreter--which respects the concrete semantics--and a family of abstract interpreters.

First we describe the parameters to the interpreter. Then we conclude the section with an implementation which is generic to these parameters.

There will be three parameters to our abstract interpreter, one of which is novel in this work:

1. The monad, novel in this work. This is the execution engine of the interpreter and captures the flow-sensitivity of the analysis.
2. The abstract domain. For our language is merely an abstraction for integers.
3. The abstraction for time. Abstract time captures the call-site sensitivity of the analysis, as introduced by [CITE].

We place each of these parameters behind an abstract interface and leave their implementations opaque for the generic monadic interpreter. We will give each of these parameters reasoning principles as we introduce them. These reasoning principles allow us to reason about the correctness of the generic interpreter independent of a particular instantiation. The goal is to factor as much of the proof-effort into what we can say about the generic interpreter. An instantiation of the interpreter need only justify that each parameter meets their local interface.

3.1 The Monad

The monad for the interpreter is capturing the *effects* of interpretation. There are two effects we wish to model in the interpreter, state and nondeterminism. The state effect will mediate how the interpreter interacts with state cells in the state space, like *Env* and *Store*. The nondeterminism effect will mediate the branching of the execution from the interpreter. Our result is that path and flow sensitivities can be recovered by altering how these effects interact in the monad.

We briefly review monad, state and nondeterminism operators and thier laws.

3.1.1 Monad Properties

To be a monad, a type operator *M* must support the *bind* operation:

$$\mathit{bind}\colon \forall\alpha\beta, M(\alpha) \to (\alpha \to M(\beta)) \to M(\beta)$$

as well as a unit for *bind* called *return*:

$$\mathit{return}\colon \forall\alpha, \alpha \to M(\alpha)$$

We use the monad laws to reason about our implementation in the absence of a particular implementatino of *bind* and *return*:

$$\mathit{bind\text{-}unit}_1\colon \mathit{bind}(\mathit{return}(a))(k) = k(a)$$

$$\mathit{bind\text{-}unit}_2\colon \mathit{bind}(m)(\mathit{return}) = m$$

$$\mathit{bind\text{-}assoc}\colon \mathit{bind}(\mathit{bind}(m)(k_1))(k_2) = \mathit{bind}(m)((a).\mathit{bind}(k_1(a))(k_2))$$

bind and *return* mean something different for each monadic effect class. For state, *bind* is a sequencer of state and *return* is the "no change in state" effect. For nondeterminism, *bind* implements a merging of multiple branches and *return* is the singleton branch. These operators capture the essence of the combination of explicit state-passing and set comprehension in the interpreter. Our interpreter will use these operators and avoid referencing an explicit configuration ς or explicit collections of results.

As is traditional with monadic programming, we use `do` and semicolon notation as syntactic sugar for *bind*. For example:

$$
\begin{aligned}
&\texttt{do} \\
&\quad a \leftarrow m \\
&\quad k(a)
\end{aligned}
$$

and

$$a \leftarrow m\ ;\ k(a)$$

are both just sugar for

$$\mathit{bind}(m)(k)$$

3.1.2 Monad State Properties

Interacting with a state component like *Env* is achieved through *get-Env* and *put-Env* effects:

$$\mathit{get\text{-}Env} \colon M(\mathit{Env})$$
$$\mathit{put\text{-}Env} \colon \mathit{Env} \to M(1)$$

We use the state monad laws to reason about state effects:

$$\mathit{put\text{-}put} \colon \mathit{put\text{-}Env}(s_1)\ ;\ \mathit{put\text{-}Env}(s_2) = \mathit{put\text{-}Env}(s_2)$$
$$\mathit{put\text{-}get} \colon \mathit{put\text{-}Env}(s)\ ;\ \mathit{get\text{-}Env} = \mathit{return}(s)$$
$$\mathit{get\text{-}put} \colon s \leftarrow \mathit{get\text{-}Env}\ ;\ \mathit{put\text{-}Env}(s) = \mathit{return}(1)$$
$$\mathit{get\text{-}get} \colon s_1 \leftarrow \mathit{get\text{-}Env}\ ;\ s_2 \leftarrow \mathit{get\text{-}Env}\ ;\ k(s_1, s_2) = s \leftarrow \mathit{get\text{-}Env}\ ;\ k(s, s)$$

The effects for *get-Store*, *get-KAddr* and *get-KStore* are identical.

3.1.3 Monad Nondeterminism Properties

Nondeterminism is achieved through operators `<0>` and `<+>` :

$$\texttt{<0>} \colon \forall \alpha, M(\alpha)$$
$$\_ \texttt{<+>} \_ \colon \forall \alpha, M(\alpha) \times M(\alpha) \to M(\alpha)$$

We use the nondeterminism laws to reason about nondeterminism effects:

$$\bot\texttt{-zero}_1 \colon \mathit{bind}(\texttt{<0>})(k) = \texttt{<0>}$$
$$\bot\texttt{-zero}_2 \colon \mathit{bind}(m)((a).\texttt{<0>}) = \texttt{<0>}$$
$$\bot\texttt{-unit}_1 \colon \texttt{<0>} \texttt{<+>}\ m = m$$
$$\bot\texttt{-unit}_2 \colon m \texttt{<+>}\ \texttt{<0>} = m$$
$$+\texttt{-assoc} \colon m_1 \texttt{<+>}\ (m_2 \texttt{<+>}\ m_3) = (m_1 \texttt{<+>}\ m_2) \texttt{<+>}\ m_3$$
$$+\texttt{-comm} \colon m_1 \texttt{<+>}\ m_2 = m_2 \texttt{<+>}\ m_1$$
$$+\texttt{-dist} \colon \mathit{bind}(m_1 \texttt{<+>}\ m_2)(k) = \mathit{bind}(m_1)(k) \texttt{<+>}\ \mathit{bind}(m_2)(k)$$

3.2 The Abstract Domain

The abstract domain is encapsulated by the *Val* type in the semantics. To parameterize over it, we leave *Val* opaque but require it support various operations. There is a constraint on *Val* its-self: it must be a join-semilattice with $\bot$ and $\sqcup$ respecting the usual laws. We require *Val* to be a join-semilattice so it can be merged in the *Store*.

The interface for integers consists of introduction and elimiation rules:

$$\mathit{int\text{-}I} \colon \mathbb{Z} \to \mathit{Val}$$
$$\mathit{int\text{-}if0\text{-}E} \colon \mathit{Val} \to \mathcal{P}(\mathit{Bool})$$

The laws for this interface are designed to induce a Galois connection between $\mathbb{Z}$ and *Val*:

$$\{true\} \sqsubseteq \mathit{int\text{-}if0\text{-}E}(\mathit{int\text{-}I}(i))\ \texttt{if}\ i = 0$$
$$\{false\} \sqsubseteq \mathit{int\text{-}if0\text{-}E}(\mathit{int\text{-}I}(i))\ \texttt{if}\ i \neq 0$$
$$v \sqsupseteq \bigsqcup_{b \in \mathit{int\text{-}if0\text{-}E}(v)} \theta(b)$$

where

$$\theta(true) = \mathit{int\text{-}I}(0)$$
$$\theta(false) = \bigsqcup_{i \in \mathbb{Z}\ |\ i \neq 0} \mathit{int\text{-}I}(i)$$

Additionally we must abstract closures:

$$\mathit{clo\text{-}I} \colon \mathit{Clo} \to \mathit{Val}$$
$$\mathit{clo\text{-}E} \colon \mathit{Val} \to \mathcal{P}(\mathit{Clo})$$

which follow similar laws:

$$\{c\} \sqsubseteq \mathit{clo\text{-}E}(\mathit{cloI}(c))$$
$$v \sqsubseteq \bigsqcup_{c \in \mathit{clo\text{-}E}(v)} \mathit{clo\text{-}I}(c)$$

The denotation for primitive operations δ must also be opaque:

$$\delta[\![\_, \_, \_]\!] \colon \mathit{IOp} \times \mathit{Val} \times \mathit{Val} \to \mathit{Val}$$

We can also give soundness laws for δ using int-I and int-if0-E:

$$\mathit{int\text{-}I}(i_1 + i_2) \sqsubseteq \delta[\![+, \mathit{int\text{-}I}(i_1), \mathit{int\text{-}I}(i_2)]\!]$$
$$\mathit{int\text{-}I}(i_1 - i_2) \sqsubseteq \delta[\![-, \mathit{int\text{-}I}(i_1), \mathit{int\text{-}I}(i_2)]\!]$$

Supporting additional primitive types like booleans, lists, or arbitrary inductive datatypes is analagous. Introduction functions inject the type into *Val*. Elimination functions project a finite set of discrete observations. Introduction and elimination operators must follow a Galois connection discipline.

Of note is our restraint from allowing operations over *Val* to have monadic effects. We set things up specifically in this way so that *Val* and the monad *M* can be varied independent of each other.

3.3 Abstract Time

The interface for abstract time is familiar from the AAM literature:

$$\mathit{tick} \colon \mathit{Exp} \times \mathit{KAddr} \times \mathit{Time} \to \mathit{Time}$$

In traditional AAM, *tick* is defined to have access to all of Σ. This comes from the generality of the framework--to account for all possibile *tick* functions. We only discuss instantiating *Addr* to support k-CFA, so we specialize the Σ parameter to $\mathit{Exp} \times \mathit{KAddr}$. Also in AAM is the opaque function $alloc \colon \mathit{Var} \times \mathit{Time} \to \mathit{Addr}$. Because we will only ever use the identity function for $alloc$, we omit its abstraction and instantiation in our development.

Remarkably, we need not state laws for *tick*. Our interpreter will always merge values which reside at the same address to achieve soundness. Therefore, any supplied implementations of *tick* is valid.

3.4 The Interpreter

We now present a generic monadic interpreter for IF paramaterized over *M*, *Val* and *Time*.

In moving our semantics to an analysis, we will need to reuse addresses in the state space. This induces *Store* and *KStore* to join when binding new values to in-use addresses. The state space for our interpreter will therefore use the following domain for *Store* and *KStore*:

$$\sigma \in \mathit{Store} : \mathit{Addr} \to \mathit{Val}$$
$$\kappa\sigma \in \mathit{KStore} : \mathit{KAddr} \to \mathcal{P}(\mathit{Frame} \times \mathit{KAddr})$$

We have already established a join-semilattice structure for *Val*. Developing a custom join-semilattice for continuations is possible, and is the key component of recent developments in pushdown abstraction. For this presentation we use $\mathcal{P}(\mathit{Frame} \times \mathit{KAddr})$ as an abstraction for continuations for simplicity.

Before defining the interpreter we define some helper functions which interact with the underlying monad *M*.

First, values in $\mathcal{P}(\alpha)$ can be lifted to monadic values $M(\alpha)$ using *return* and <0>, which we name $\uparrow_p$:

$$\uparrow_p : \forall\alpha, \mathcal{P}(\alpha) \to M(\alpha)$$
$$\uparrow_p (\{a_1..a_n\}) \coloneqq \mathit{return}(a_1) \texttt{ <+> .. <+> } \mathit{return}(a_n)$$

Allocating addresses and updating time can be implemented using monadic state effects:

$$\begin{aligned}
&\mathit{allocM} : \mathit{Var} \to M(\mathit{Addr}) \\
&\mathit{allocM}(x) \coloneqq \texttt{do} \\
&\quad \tau \leftarrow \mathit{get\text{-}Time} \\
&\quad \mathit{return}(x, \tau) \\
&\kappa\mathit{allocM} : M(\mathit{KAddr}) \\
&\kappa\mathit{allocM} \coloneqq \texttt{do} \\
&\quad \tau \leftarrow \mathit{get\text{-}Time} \\
&\quad \mathit{return}(\tau) \\
&\mathit{tickM} : \mathit{Exp} \to M(1) \\
&\mathit{tickM}(e) = \texttt{do} \\
&\quad \tau \leftarrow \mathit{get\text{-}Time} \\
&\quad \kappa l \leftarrow \mathit{get\text{-}KAddr} \\
&\quad \mathit{put\text{-}Time}(\mathit{tick}(e, \kappa l, \tau))
\end{aligned}$$

Finally, we introduce helper functions for manipulating stack frames:

$$\begin{aligned}
&\mathit{push} : \mathit{Frame} \to M(1) \\
&\mathit{push}(fr) \coloneqq \texttt{do} \\
&\quad \kappa l \leftarrow \mathit{get\text{-}KAddr} \\
&\quad \kappa\sigma \leftarrow \mathit{get\text{-}KStore} \\
&\quad \kappa l' \leftarrow \kappa\mathit{allocM} \\
&\quad \mathit{put\text{-}KStore}(\kappa\sigma \sqcup [\kappa l' \mapsto \{fr :: \kappa l\}]) \\
&\quad \mathit{put\text{-}KAddr}(\kappa l') \\
&\mathit{pop} : M(\mathit{Frame}) \\
&\mathit{pop} \coloneqq \texttt{do} \\
&\quad \kappa l \leftarrow \mathit{get\text{-}KAddr} \\
&\quad \kappa\sigma \leftarrow \mathit{get\text{-}KStore} \\
&\quad fr :: \kappa l' \leftarrow \uparrow_p (\kappa\sigma(\kappa l)) \\
&\quad \mathit{put\text{-}KAddr}(\kappa l') \\
&\quad \mathit{return}(fr)
\end{aligned}$$

To implement our interpreter we define a denotation function for atomic expressions and a step function for compound expressions. The denotation for atomic expressions is written as a monadic computation from atomic expresssions to values.

$$\begin{aligned}
&A[\![_]\!] \in \mathit{Atom} \to M(\mathit{Val}) \\
&A[\![i]\!] \coloneqq \mathit{return}(\mathit{int\text{-}I}(i)) \\
&A[\![x]\!] \coloneqq \texttt{do} \\
&\quad \rho \leftarrow \mathit{get\text{-}Env} \\
&\quad \sigma \leftarrow \mathit{get\text{-}Store} \\
&\quad l \leftarrow \uparrow_p (\rho(x)) \\
&\quad \mathit{return}(\sigma(x)) \\
&A[\![\underline{\boldsymbol{\lambda}}(x).e]\!] \coloneqq \texttt{do} \\
&\quad \rho \leftarrow \mathit{get\text{-}Env} \\
&\quad \mathit{return}(\mathit{clo\text{-}I}(\langle \underline{\boldsymbol{\lambda}}(x).e, \rho \rangle))
\end{aligned}$$

The step function is written as a monadic computation from expressions to the next expression to evaluate, in small step style. The definition for operators is simple: it merely pushes a stack from and returns the first operand:

$$\begin{aligned}
&\mathit{step} : \mathit{Exp} \to M(\mathit{Exp}) \\
&\mathit{step}(e_1 \odot e_2) \coloneqq \texttt{do} \\
&\quad \mathit{tickM}(e_1 \odot e_2) \\
&\quad \mathit{push}(\langle \square \odot e_2 \rangle) \\
&\quad \mathit{return}(e_1)
\end{aligned}$$

The definition for atomic expressions must pop and inspect the stack and perform the denotation of the operation:

$$
\begin{array}{l}
\mathit{step}(a) \coloneqq \texttt{do} \\
\quad \mathit{tickM}(a) \\
\quad \mathit{fr} \leftarrow \mathit{pop} \\
\quad v \leftarrow A[\![a]\!] \\
\quad \texttt{case}\ \ \mathit{fr}\ \ \texttt{of} \\
\qquad \langle \square \odot e \rangle \rightarrow \texttt{do} \\
\qquad\quad \mathit{push}(\langle v \odot \square \rangle) \\
\qquad\quad \mathit{return}(e) \\
\qquad \langle v' \ @ \square \rangle \rightarrow \texttt{do} \\
\qquad\quad \langle \underline{\boldsymbol{\lambda}}(x).e, \rho' \rangle \leftarrow\!\uparrow_p (\mathit{clo\text{-}E}(v')) \\
\qquad\quad l \leftarrow \mathit{alloc}(x) \\
\qquad\quad \sigma \leftarrow \mathit{get\text{-}Store} \\
\qquad\quad \mathit{put\text{-}Env}(\rho'[x \mapsto l]) \\
\qquad\quad \mathit{put\text{-}Store}(\sigma[l \mapsto v]) \\
\qquad\quad \mathit{return}(e) \\
\qquad \langle v' \oplus \square \rangle \rightarrow \texttt{do} \\
\qquad\quad \mathit{return}(\delta(\oplus, v', v)) \\
\qquad \langle \underline{\textbf{if0}}(\square)\{e_1\}\{e_2\} \rangle \rightarrow \texttt{do} \\
\qquad\quad b \leftarrow\!\uparrow_p (\mathit{int\text{-}if0\text{-}E}(v)) \\
\qquad\quad \texttt{if}(b)\ \ \texttt{then}\ \ \mathit{return}(e_1)\ \ \texttt{else}\ \ \mathit{return}(e_2)
\end{array}
$$

We can also implement abstract garbage collection in a fully general away against the monadic effect interface:

$$
\begin{array}{l}
\mathit{gc} \colon \mathit{Exp} \rightarrow M(1) \\
\mathit{gc}(e) \coloneqq \texttt{do} \\
\quad \rho \leftarrow \mathit{get\text{-}Env} \\
\quad \sigma \leftarrow \mathit{get\text{-}Store} \\
\quad \kappa\sigma \leftarrow \mathit{get\text{-}KStore} \\
\quad l*_0 \leftarrow R_0(\rho, e) \\
\quad \kappa l_0 \leftarrow \mathit{get\text{-}KAddr} \\
\quad \text{let}\ \ l*' \coloneqq \mu(\theta).l *_0 \cup \theta \cup R[\sigma](\theta) \\
\quad \text{let}\ \ \kappa l*' \coloneqq \mu(\kappa\theta).\{\kappa l_0\} \cup \kappa\theta \cup KR[\kappa\sigma](\kappa\theta) \\
\quad \mathit{put\text{-}Store}(\{l \mapsto \sigma(l) \mid l \in l*'\}) \\
\quad \mathit{put\text{-}KStore}(\{\kappa l \mapsto \kappa\sigma(\kappa l) \mid \kappa l \in \kappa l*'\})
\end{array}
$$

where R_0 is defined as before and R, KR and $R - Clo$ are defined:

$$
\begin{array}{l}
R \colon \mathit{Store} \rightarrow \mathcal{P}(\mathit{Addr}) \rightarrow \mathcal{P}(\mathit{Addr}) \\
R[\sigma](\theta) \coloneqq \{l' \mid l' \in R - \mathit{Clo}(c)\ ;\ c \in \mathit{clo\text{-}E}(v)\ ;\ v \in \sigma(l)\ ;\ l \in \theta\} \\
R - \mathit{Clo} \colon \mathit{Clo} \rightarrow \mathcal{P}(\mathit{Addr}) \\
R - \mathit{Clo}(\langle \underline{\boldsymbol{\lambda}}(x).e, \rho \rangle) \coloneqq \{\rho(x) \mid x \in \mathit{FV}(\underline{\boldsymbol{\lambda}}(x).e)\} \\
KR \colon \mathit{KStore} \rightarrow \mathcal{P}(\mathit{KAddr}) \rightarrow \mathcal{P}(\mathit{KAddr}) \\
KR[\sigma](\kappa\theta) \coloneqq \{\pi_2(\mathit{fr}) \mid \mathit{fr} \in \kappa\sigma(\kappa l)\ ;\ \kappa l \in \theta\}
\end{array}
$$

To execute the interpreter we must introduce one more parameter. In the concrete semantics, execution takes the form of a least-fixed-point computation over the collecting semantics This in general requires a join-semilattice structure for some Σ and a transition function $\Sigma \rightarrow \Sigma$. We bridge this gap between monadic interpreters and transition functions with an extra constraint on the monad M. We require that monadic actions $\alpha \rightarrow M(\beta)$ form a Galois connection with a transition system $\Sigma \rightarrow \Sigma$.

There is one last parameter to our development: a connection between our monadic interpreter and a state space transition system. We state this connection formally as a Galois connection $(\Sigma \rightarrow \Sigma) \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} (\mathit{Exp} \rightarrow M(\mathit{Exp}))$. This Galois connection serves two purposes. First, it allows us to implement the analysis by converting our interpreter to the transition system $\Sigma \rightarrow \Sigma$ through γ. Second, this Galois connection serves to *transport other Galois connections*. For example, given concrete and abstract versions of *Val*, we carry $\mathbf{Val} \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \widehat{\mathbf{Val}}$ through the Galois connection to establish $\boldsymbol{\Sigma} \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \widehat{\boldsymbol{\Sigma}}$.

A collecting-semantics execution of our interpreter is defined as:

$$\mu(X).\varsigma_0 \sqcup X \sqcup \gamma(\mathit{step})(X)$$

where ς_0 is the injection of the initial program e_0 into Σ.

4. Recovering Interpreters

4.1 Recovering a Concrete Interpreter

For the concrete value space we instantiate $\mathbf{Val}$ to a powerset of *Val*.

$$v \in \mathbf{Val} \coloneqq \mathcal{P}(\mathit{Val})$$

The concrete value space $\mathbf{Val}$ has straightforward introduction and elimination rules:

$$
\begin{array}{l}
\mathit{int\text{-}I} \colon \mathbb{Z} \rightarrow \mathbf{Val} \\
\mathit{int\text{-}I}(i) \coloneqq \{i\} \\
\mathit{int\text{-}if0\text{-}E} \colon \mathbf{Val} \rightarrow \mathcal{P}(\mathit{Bool}) \\
\mathit{int\text{-}if0\text{-}E}(v) \coloneqq \{\mathit{true} \mid 0 \in v\} \cup \{\mathit{false} \mid i \in v \wedge i \neq 0\}
\end{array}
$$

and the concrete δ you would expect:

$$
\begin{array}{l}
\delta[\![\_, \_, \_]\!] \colon \mathit{IOp} \times \mathbf{Val} \times \mathbf{Val} \rightarrow \mathbf{Val} \\
\delta[\![+, v_1, v_2]\!] \coloneqq \{i_1 + i_2 \mid i_1 \in v_1\ ;\ i_2 \in v_2\} \\
\delta[\![-, v_1, v_2]\!] \coloneqq \{i_1 - i_2 \mid i_1 \in v_1\ ;\ i_2 \in v_2\}
\end{array}
$$

Proposition 1. $\mathbf{Val}$ *satisfies the abstract domain laws from section X.*

Concrete time $\mathbf{Time}$ captures program contours as a product of *Exp* and *KAddr*:

$$\tau \in \mathbf{Time} \coloneqq (\mathit{Exp} \times \mathit{KAddr})*$$

and *tick* is just a cons operator:

$$
\begin{array}{l}
\mathit{tick} \colon \mathit{Exp} \times \mathit{KAddr} \times \mathit{Time} \rightarrow \mathit{Time} \\
\mathit{tick}(e, \kappa l, \tau) \coloneqq (e, \kappa l) :: \tau
\end{array}
$$

For the concrete monad we instantiate M to a path-sensitive $\mathbf{M}$ which contains a powerset of concrete state space components.

$$\psi \in \Psi^{cm} \coloneqq \mathit{Env} \times \mathit{Store} \times \mathit{KAddr} \times \mathit{KStore} \times \mathit{Time}$$
$$m \in \mathbf{M}(\alpha) \coloneqq \Psi^{cm} \to \mathcal{P}(\alpha \times \Psi^{cm})$$

Monadic operators bind^{cm} and return^{cm} encapsulate both state-passing and set-flattening:

$$\mathit{bind}^{cm} \colon \forall \alpha, \mathbf{M}(\alpha) \to (\alpha \to \mathbf{M}(\beta)) \to \mathbf{M}(\beta)$$
$$\mathit{bind}^{cm}(m)(f)(\psi) \coloneqq \{(y, \psi'') \mid (y, \psi'') \in f(a)(\psi') \,;\, (a, \psi') \in m(\psi)\}$$
$$\mathit{return}^{cm} \colon \forall \alpha, \alpha \to \mathbf{M}(\alpha)$$
$$\mathit{return}^{cm}(a)(\psi) \coloneqq \{(a, \psi)\}$$

State effects merely return singleton sets:

$$\mathit{get\text{-}Env}^{cm} \colon \mathbf{M}(\mathit{Env})$$
$$\mathit{get\text{-}Env}^{cm}(\langle \rho, \sigma, \kappa, \tau \rangle) \coloneqq \{(\rho, \langle \rho, \sigma, \kappa, \tau \rangle)\}$$
$$\mathit{put\text{-}Env}^{cm} \colon \mathit{Env} \to \mathcal{P}(1)$$
$$\mathit{put\text{-}Env}^{cm}(\rho')(\langle \rho, \sigma, \kappa, \tau \rangle) \coloneqq \{(1, \langle \rho', \sigma, \kappa, \tau \rangle)\}$$

Nondeterminism effects are implemented with set union:

$$<0>^{cm} \colon \forall \alpha, \mathbf{M}(\alpha)$$
$$<0>^{cm}(\psi) \coloneqq \{\}$$
$$\_ <+>^{cm} \_ \colon \forall \alpha, \mathbf{M}(\alpha) \times \mathbf{M}(\alpha) \to \mathbf{M}(\alpha)$$
$$(m_1 <+>^{cm} m_2)(\psi) \coloneqq m_1(\psi) \cup m_2(\psi)$$

Proposition 2. $\mathbf{M}$ *satisfies monad, state, and nondeterminism laws.*

Finally, we must establish a Galois connection between $\mathit{Exp} \to \mathbf{M}(\mathit{Exp})$ and $\mathbf{\Sigma} \to \mathbf{\Sigma}$ for some choice of $\mathbf{\Sigma}$. For the path sensitive monad $\mathbf{M}$ instantiate with $\mathbf{Val}$ and $\mathbf{Time}$, , $\mathbf{\Sigma}$ is defined:

$$\mathbf{\Sigma} \coloneqq \mathcal{P}(\mathit{Exp} \times \Psi^{cm})$$

The Galois connection between $\mathbf{M}$ and $\mathbf{\Sigma}$ is straightforward:

$$\gamma^{cm} \colon (\mathit{Exp} \to \mathbf{M}(\mathit{Exp})) \to \mathbf{\Sigma} \to \mathbf{\Sigma}$$
$$\gamma^{cm}(f)(e\psi*) \coloneqq \{(e, \psi') \mid (e, \psi') \in f(e)(\psi) \,;\, (e, \psi) \in e\psi*\}$$
$$\alpha^{cm} \colon (\mathbf{\Sigma} \to \mathbf{\Sigma}) \to \mathit{Exp} \to \mathbf{M}(\mathit{Exp})$$
$$\alpha^{cm}(f)(e)(\psi) \coloneqq f(\{(e, \psi)\})$$

The injection ς_0^{cm} for a program e_0 is:

$$\varsigma_0^{cm} \coloneqq \{\langle e, \bot, \bot, , \bot, \rangle\}$$

Proposition 3. γ^{cm} *and* α^{cm} *form an isomorphism.*

Corollary 1. γ^{cm} *and* α^{cm} *form a Galois connection.*

4.2 Recovering an Abstract Interpreter

To arrive at an abstract interpreter we need seek only seek a monad AM that has a Galois connection to a finite state space $\widehat{\mathbf{\Sigma}}$.

We pick a simple abstraction for integers, $\{-, 0, +\}$, although our technique scales seamlessly to other domains. As a consequence, the value type $\widehat{\mathbf{Val}}$ turns into a powerset of abstract values:

$$\widehat{\mathbf{Val}} \coloneqq \mathcal{P}(\mathit{Clo} + \{-, 0, +\})$$

Introduction and elimination functions for $\widehat{\mathbf{Val}}$ are defined:

$$\mathit{int\text{-}I} \colon \mathbb{Z} \to \widehat{\mathbf{Val}}$$
$$\mathit{int\text{-}I}(i) \coloneqq$$
$$\quad - \texttt{ if } i < 0$$
$$\quad 0 \texttt{ if } i = 0$$
$$\quad + \texttt{ if } i > 0$$
$$\mathit{int\text{-}if0\text{-}E} \colon \widehat{\mathbf{Val}} \to \mathcal{P}(\mathit{Bool})$$
$$\mathit{int\text{-}if0\text{-}E}(v) \coloneqq \{\mathit{true} \mid 0 \in v\} \cup \{\mathit{false} \mid - \in v \vee + \in v\}$$

Introduction and elmination for *Clo* is identical to the concrete domain.

The abstract $\widehat{\delta}$ operator is defined:

$$\widehat{\delta} \colon \mathit{IOp} \times \widehat{\mathbf{Val}} \times \widehat{\mathbf{Val}} \to \widehat{\mathbf{Val}}$$
$$\begin{aligned}\widehat{\delta}(+, v_1, v_2) \coloneqq \{i \mid 0 \in v_1 \wedge i \in v_2\} \\ \cup \{i \mid i \in v_1 \wedge 0 \in v_2\} \\ \cup \{+ \mid + \in v_1 \wedge + \in v_2\} \\ \cup \{- \mid - \in v_1 \wedge - \in v_2\} \\ \cup \{-, 0, + \mid + \in v_1 \wedge - \in v_2\} \\ \cup \{-, 0, + \mid - \in v_1 \wedge + \in v_2\}\end{aligned}$$

The definition for $\widehat{\delta}(-, v_1, v_2)$ is analagous.

Next we abstract *Time* to the finite domain of k-truncated lists of execution contexts:

$$\mathit{Time} \coloneqq (\mathit{Exp} \times \mathit{KAddr}) *$$

The *tick* operator becomes cons followed by k-truncation:

$$\mathit{tick} \colon \mathit{Exp} \times \mathit{KAddr} \times \mathit{Time} \to \mathit{Time}$$
$$\mathit{tick}(e, \kappa l, \tau) = \lfloor (e, \kappa l) :: \tau \rfloor$$

The monad AM need not change in implementation from $\mathbf{M}$; they are identical up to choices for $AStore$ (which maps to $\widehat{\mathbf{Val}}$) and $\widehat{\mathbf{Time}}$.

$$\psi \in \Psi \coloneqq \mathit{Env} \times \mathit{AStore} \times \mathit{KAddr} \times \mathit{KStore} \times \widehat{\mathbf{Time}}$$

The resulting state space $\widehat{\mathbf{\Sigma}}$ is finite, and its least-fixed-point iteration will give a sound and computable analysis.

5. Varying Path and Flow Sensitivity

We are able to recover a flow-insensitivity in the analysis through a new definition for M: AM^{fi}. To do this we pull *Store* out of the powerset and exploit its join-semilattice structure:

$$\Psi^{fi} \coloneqq \mathit{Env} \times \mathit{KAddr} \times \mathit{KStore} \times \mathit{Time}$$
$$AM^{fi}(\alpha) \coloneqq \Psi^{fi} \times \mathit{Store} \times \mathcal{P}(\alpha \times \Psi^{fi}) \times \mathit{Store}$$

The monad operator bind^{fi} performs the store merging needed to capture a flow-insensitive analysis.

$$\mathit{bind}^{fi} \colon \forall \alpha\beta, AM^{fi}(\alpha) \to (\alpha \to AM^{fi}(\beta)) \to AM^{fi}(\beta)$$
$$\mathit{bind}^{fi}(m)(f)(\psi, \sigma) \coloneqq (\{bs_{11}..bs_{n1}..bs_{nm}\}, \sigma_1 \sqcup .. \sqcup \sigma_n)$$
where
$$(\{(a_1, \psi_1)..(a_n, \psi_n)\}, \sigma') \coloneqq m(\psi, \sigma)$$
$$(\{b\psi_{i1}..b\psi_{im}\}, \sigma_i) \coloneqq f(a_i)(\psi_i, \sigma')$$

The unit for bind^{fi} returns one nondeterminism branch and a single store:

$$\mathit{return}^{fi} \colon \forall \alpha, \alpha \to AM^{fi}(\alpha)$$
$$\mathit{return}^{fi}(a)(\psi, \sigma) \coloneqq (\{a, \psi\}, \sigma)$$

State effects *get-Env* and *put-Env* are also straightforward, returning one branch of nondeterminism:

$$\mathit{get\text{-}Env}^{fi} \colon AM^{fi}(\mathit{Env})$$
$$\mathit{get\text{-}Env}^{fi}(\langle \rho, \kappa, \tau \rangle, \sigma) \coloneqq (\{(\rho, \langle \rho, \kappa, \tau \rangle)\}, \sigma)$$
$$\mathit{put\text{-}Env}^{fi} \colon \mathit{Env} \to AM^{fi}(1)$$
$$\mathit{put\text{-}Env}^{fi}(\rho')(\langle \rho, \kappa, \tau \rangle, \sigma) \coloneqq (\{(1, \langle \rho', \kappa, \tau \rangle)\}, \sigma)$$

State effects *get-Store* and *put-Store* are analagous to *get-Env* and *put-Env*:

$$\mathit{get\text{-}Store}^{fi} \colon AM^{fi}(\mathit{Env})$$
$$\mathit{get\text{-}Store}^{fi}(\langle \rho, \kappa, \tau \rangle, \sigma) \coloneqq (\{(\sigma, \langle \rho, \kappa, \tau \rangle\}, \sigma)$$
$$\mathit{put\text{-}Store}^{fi} \colon \mathit{Store} \to AM^{fi}(1)$$
$$\mathit{put\text{-}Store}^{fi}(\sigma')(\langle \rho, \kappa, \tau \rangle, \sigma) \coloneqq (\{(1, \langle \rho, \kappa, \tau \rangle)\}, \sigma')$$

Nondeterminism operations union the powerset and join the store pairwise:

$$\texttt{<0>}^{fi} \colon \forall \alpha, M(\alpha)$$
$$\texttt{<0>}^{fi}(\psi, \sigma) \coloneqq (\{\}, \bot)$$
$$\_ \texttt{<+>} \_ \colon \forall \alpha, M(\alpha) \times M(\alpha) \to M\alpha$$
$$(m_1 \texttt{<+>} m_2)(\psi, \sigma) \coloneqq (a\psi *_1 \cup a\psi *_2, \sigma_1 \sqcup \sigma_2)$$
where $(a\psi *_i, \sigma_i) \coloneqq m_i(\psi, \sigma)$

Finally, the Galois connection relating AM^{fi} to a state space transition over $\widehat{\mathbf{\Sigma}}^{fi}$ must also compute nondeterminism unions and store joins:

$$\widehat{\mathbf{\Sigma}}^{fi} \coloneqq \mathcal{P}(\mathit{Exp} \times \Psi^{fi}) \times \mathit{Store}$$
$$\gamma^{fi} \colon (\mathit{Exp} \to AM^{fi}(\mathit{Exp})) \to (\Sigma^{fi} \to \Sigma^{fi})$$
$$\gamma^{fi}(f)(e\psi *, \sigma) \coloneqq (\{e\psi_{11}..e\psi_{n1}..e\psi_{nm}\}, \sigma_1 \sqcup .. \sqcup \sigma_n)$$
where
$$\{(e_1, \psi_1)..(e_n, \psi_n)\} \coloneqq e\psi *$$
$$(\{e\psi_{i1}..e\psi_{im}\}, \sigma_i) \coloneqq f(e_i)(\psi_i, \sigma)$$
$$\alpha^{fi} \colon (\Sigma^{fi} \to \Sigma^{fi}) \to (\mathit{Exp} \to AM^{fi}(\mathit{Exp}))$$
$$\alpha^{fi}(f)(e)(\psi, \sigma) \coloneqq f(\{(e, \psi)\}, \sigma)$$

Proposition 4. *γ^{fi} and α^{fi} form an isomorphism.*

Corollary 2. *γ^{fi} and α^{fi} form a Galois connection.*

Proposition 5. *There exists Galois connection $\mathbf{\Sigma}\alpha_1\gamma_1\widehat{\mathbf{\Sigma}}\alpha_2\gamma_2\widehat{\mathbf{\Sigma}}^{fi}$ and $\alpha_1 C\gamma(\mathit{step})\gamma_1 \sqsubseteq A\gamma(\mathit{step}) \sqsubseteq \gamma_2 A\gamma^{fi}(\mathit{step})\alpha_2$*

The first Galois connection $\mathbf{\Sigma}\alpha_1\gamma_1\widehat{\mathbf{\Sigma}}$ is justified by the Galois connections between $\mathbf{Val} \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \widehat{\mathbf{Val}}$ and $\mathbf{Time} \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \widehat{\mathbf{Time}}$. The second Galois connection $\widehat{\mathbf{\Sigma}}\alpha_2\gamma_2\widehat{\mathbf{\Sigma}}^{fi}$ is justified by first calculating the Galois connection between monads AM and $\mathbf{M}$, and then transporting it through their respective Galois connections to $\widehat{\mathbf{\Sigma}}$ and $\widehat{\mathbf{\Sigma}}^{fi}$. These proofs are tedious calculations over the definitions which we do not repeat here. However, we will recover these proof in a later section through our compositional framework which greatly reduces the proof burden.

We note that the implementation for our interpreter and abstract garbage collector remain the same. They both scale seamlessly to flow-sensitive and flow-insensitive variants when instantiated with the appropriate monad.

6. A Compositional Monadic Framework

In our framework thus far, any modification to the interpreter requires redesigning the monad M and constructing new proofs. We want to avoid reconstructing complicated monads for our interpreters, especially as languages and analyses grow and change. Even more, we want to avoid reconstructing the *proofs* that these changes will necessarily alter. Toward this goal we introduce a compositional framework for constructing monads which are correct-by-construction using a restricted class of monad transformer.

There are two types of monadic effects used in the monadic interprer: state and nondeterminism. There is a monad transformer for adding state effects to existing monads, called the state monad tranformer:

$$S[\_] \colon (\mathit{Type} \to \mathit{Type}) \to (\mathit{Type} \to \mathit{Type})$$
$$S[s](m)(\alpha) \coloneqq s \to m(\alpha \times s)$$

Monadic actions *bind* and *return* (and their laws) use the underlying monad:

$$
\begin{aligned}
&\mathit{bind}\colon \forall \alpha\beta, S[s](m)(\alpha) \to (\alpha \to S[s](m)(\beta)) \to S[s](m)(\beta)\\
&\mathit{bind}(m)(f)(s) \coloneqq \texttt{do}\\
&\quad (x, s') \leftarrow_m m(s)\\
&\quad f(x)(s')\\
&\mathit{return}\colon \forall \alpha m, \alpha \to S[s](m)(\alpha)\\
&\mathit{return}(x)(s) \coloneqq \mathit{return}_m(x, s)
\end{aligned}
$$

State actions get and put expose the cell of state while interacting with the underlying monad m:

$$
\begin{aligned}
&get\colon S[s](m)(s)\\
&get(s) \coloneqq \mathit{return}_m(s, s)\\
&put\colon s \to S[s](m)(1)\\
&put(s')(s) \coloneqq \mathit{return}_m(1, s')
\end{aligned}
$$

and the state monad transformer is able to transport nondeterminism effects from the underlying monad:

$$
\begin{aligned}
&\texttt{<0>}\colon \forall \alpha, S[s](m)(\alpha)\\
&\texttt{<0>}(s) \coloneqq \texttt{<0>}_m\\
&\_ \texttt{<+>} \_\colon \forall \alpha, S[s](m)(\alpha) x S[s](m)(\alpha) \to S[s](m)(\alpha)\\
&(m_1 \texttt{<+>} m_2)(s) \coloneqq m_1(s) \quad \texttt{<+>} \quad _m\, m_2(s)
\end{aligned}
$$

The state monad transformer was introduced by Mark P. Jones in [X].

We develop a new monad transformer for nondeterminism which can compose with state in both directions.

$$
\begin{aligned}
&\mathcal{P}\colon (Type \to Type) \to (Type \to Type)\\
&\mathcal{P}(m)(\alpha) \coloneqq m(\mathcal{P}(\alpha))
\end{aligned}
$$

Monadic actions *bind* and *return* require that the underlying monad be a join-semilattice functor:

$$
\begin{aligned}
&\mathit{bind}_p\colon \forall \alpha\beta, \mathcal{P}(m)(\alpha) \to (\alpha \to \mathcal{P}(m)(\beta)) \to \mathcal{P}(m)(\beta)\\
&\mathit{bind}_p(m)(f) \coloneqq \texttt{do}\\
&\quad \{x_1..x_n\} \leftarrow_m m\\
&\quad f(x_1) \sqcup_m .. \sqcup_m f(x_n)\\
&\mathit{return}_p\colon \forall \alpha, \alpha \to \mathcal{P}(m)(\alpha)\\
&\mathit{return}_p(x) \coloneqq \mathit{return}_m(\{x\})
\end{aligned}
$$

Nondterminism actions $\texttt{<0>} and \texttt{<+>}$ interact with the join-semilattice functorality of the underlying monad m:

$$
\begin{aligned}
&\texttt{<0>}_p\colon \forall \alpha, \mathcal{P}(m)(\alpha)\\
&\texttt{<0>}_p \coloneqq \bot\\
&\_ \texttt{<+>}\ _p \_\colon \forall \alpha, \mathcal{P}(m)(\alpha) x \mathcal{P}(m)(\alpha) \to \mathcal{P}(m)(\alpha)\\
&m_1 \texttt{<+>}\ _p m_2 \coloneqq m_1 \sqcup_m m_2
\end{aligned}
$$

and the nondeterminism monad transformer is able to transport state effects from the underlying monad:

$$
\begin{aligned}
&get_p\colon \mathcal{P}(m)(s)\\
&get_p = map_m((s).\{s\})(get_m)\\
&put_p\colon s \to \mathcal{P}(m)(s)\\
&put_p(s) = map_m((1).\{1\})(put_m(s))
\end{aligned}
$$

Proposition: $\mathcal{P}$ is a transformer for monads which are also join semi-lattice functors.

Our correctness framework requires that monadic actions in *M* map to state space transitions in Σ. We establish this property in addition to monadic actions and effects for state and nondeterminism monad transformers. We call this property $MonadStep$, where monadic acations in *M* admit a Galois connection to transitions in Σ:

$$
mstep\colon \forall \alpha\beta, (\alpha \to \boldsymbol{M}(\beta)) \xleftarrow[\alpha]{\gamma} (\Sigma(\alpha) \to \Sigma(\beta))
$$

We now show that the monad transformers for state and nondeterminism transport this property in addition to monadic operations.

For the state monad transformer $S[s]$ mstep is defined:

$$
\begin{aligned}
&mstep - \gamma\colon \forall \alpha\beta m, (\alpha \to S[s](m)(\beta)) \to (\Sigma_m(\alpha \times s) \to \Sigma_m(\beta \times s))\\
&mstep - \gamma(f) \coloneqq mstep_m - \gamma((a, s).f(a)(s))
\end{aligned}
$$

For the nondeterminism transformer $\mathcal{P}$, mstep has two possible definitions. One where Σ is ΣP:

$$
\begin{aligned}
&mstep_{p1} - \gamma\colon \forall \alpha\beta m, (\alpha \to \mathcal{P}(m)(\beta)) \to (\Sigma_m(\mathcal{P}(\alpha)) \to \Sigma_m(\mathcal{P}(\beta)))\\
&mstep_{p1} - \gamma(f) \coloneqq mstep_m - \gamma((\{x_1..x_n\}).f(x_1) \texttt{<+>} .. \texttt{<+>} f(x_n))
\end{aligned}
$$

and one where Σ is $P\Sigma$:

$$
\begin{aligned}
&mstep_{p2} - \gamma\colon \forall \alpha\beta m, (\alpha \to \mathcal{P}(m)(\beta)) \to (\mathcal{P}(\Sigma_m(\alpha)) \to \mathcal{P}(\Sigma_m(\beta)))\\
&mstep_{p2} - \gamma(f)(\{\varsigma_1..\varsigma_n\}) \coloneqq a\Sigma P_1 \cup .. \cup a\Sigma P_n\\
&\quad \text{where}\\
&\qquad commuteP\colon \forall \alpha, \Sigma_m(\mathcal{P}(\alpha)) \to \mathcal{P}(\Sigma_m(\alpha))\\
&\qquad a\Sigma P_i \coloneqq commuteP - \gamma(mstep_m - \gamma(f)(\varsigma_i))
\end{aligned}
$$

The operation $computeP$ must be defined for the underlying Σ. This property is true for the identiy monad, and is preserved by $S[s]$ when Σ is also a functor:

$$
\begin{aligned}
&commuteP - \gamma\colon \forall \alpha, \Sigma_m(\mathcal{P}(\alpha) \times s) \to \mathcal{P}(\Sigma_m(\alpha \times s))\\
&commuteP - \gamma \coloneqq commuteP_m map((\{\alpha_1..\alpha_n\}, s).\{(\alpha_1, s)..(\alpha_n, s)\})
\end{aligned}
$$

The γ side of commuteP is the only Galois connection mapping that loses information in the α direction. Therefore, $mstep$ and $mstep_{p1}$ are really isomorphism transformers, and $mstep_{p2}$ is the only Galois connection transformer.

[QUESTION: should I give the definitions for the α maps here? -DD]

For convenience, we name the pairing of $\mathcal{P}$ with $mstep_1$ FI, and with $mstep_{p2}$ FS for flow insensitive and flow sensitive respectively.

We can now build monad transformer stacks from combinations of $S[s]$, FI and FS that have the following properties:

- The resulting monad has the combined effects of all pieces of the transformer stack.
- Actions in the resulting monad map to a state space transition system $\Sigma \to \Sigma$ for some Σ.
- Galois connections between states s_1 and s_2 are transported along the Galois connection between $(\alpha \to S[s_1](m)(\beta)) \underset{\gamma}{\overset{\alpha}{\leftrightarrows}} (\Sigma[s_1](\alpha) \to \Sigma[s_1](\beta))$ and $(\alpha \to S[s_2](m)(\beta)) \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} (\Sigma[s_2](\alpha) \to \Sigma[s_2](\beta))$ resulting in $(\Sigma[s_1](\alpha) \to \Sigma[s_1](\beta))\alpha\beta(\Sigma[s_2](\alpha) \to \Sigma[s_2](\beta))$.

We can now instantiate our interpreter to the following monad stacks.

- S[*Env*]S[*Store*]S[*KAddr*]S[*KStore*]S[*Time*]FS
  - This yields a path-sensitive flow-sensitive analysis.
- S[*Env*]S[*KAddr*]S[*KStore*]S[*Time*]FSS[*Store*]
  - This yeilds a path-insensitive flow-sensitive analysis.
- S[*Env*]S[*KAddr*]S[*KStore*]S[*Time*]FIS[*Store*]
  - This yields a path-insensitive flow-insensitive analysis.

Furthermore, the final Galois connection for each state space is justified from individual Galois connections between state space components.