# 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 a family of an abstract interpreter. Section X gives the concrete semantics for a small functional language. Section X shows the full definition of a concrete monadic interpreter. Section X shows our compositional metatheory framework built on monad transformers.

## 2. Semantics

Our language of study is `λIF`:

```
i ∈ ℤ
x ∈ Var
a ∈ Atom ::= i | x | λ(x).e
iop ∈ IOp ::= + | -
op ∈ Op ::= iop | @
e∈ Exp ::= a | e op e | if0(e){e}{e}
```

(The operator `@` is syntax for function application. We define `op` as a single syntactic class for all operators to simplify presentation.) We begin with a concrete semantics for `λIF` which makes allocation explicit. Allocation is made explicit to make the semantics more amenable to abstraction and abstract garbage collection.

The concrete semantics for `λIF`:

```
τ ∈ Time := ℤ

l ∈ Addr := Var × Time
ρ ∈ Env := Var ⇀ Addr
σ ∈ Store := Addr ⇀ Val
c ∈ Clo ::= ⟨λ(x).e,ρ⟩
v ∈ Val ::= i | c

κl ∈ KAddr := Time
κσ ∈ KStore := KAddr ⇀ Frame × KAddr
fr ∈ Frame ::= ⟨□ op e⟩ | ⟨v op □⟩ | ⟨if0(□){e}{e}⟩

ς ∈ Σ ::= Exp × Env × Store × KAddr × KStore

A⟦_,_,_⟧ ∈ Env × Store × Atom ⇀ Val
A⟦ρ,σ,i⟧ := i
A⟦ρ,σ,x⟧ := σ(ρ(x))
A⟦ρ,σ,λ(x).e⟧ := ⟨λ(x).e,ρ⟩

δ⟦_,_,_⟧ ∈ IOp × ℤ × ℤ → ℤ
δ⟦+,i₁,i₂⟧ := i₁ + i₂
δ⟦-,i₁,i₂⟧ := i₁ - i₂

_-->_ ∈ P(Σ × Σ)
⟨e₁ op e₂,ρ,σ,κl,κσ,τ⟩ --> ⟨e₁,ρ,σ,τ,κσ',τ+1⟩
  where κσ' := κσ[τ ↦ ⟨□ op e₂⟩∷κl]
⟨a,ρ,σ,κl,κσ,τ⟩ --> ⟨e,ρ,σ,τ,κσ',tick(τ)⟩
  where ⟨□ op e⟩∷κl' := κσ(κl)
        κσ' := κσ[τ ↦ ⟨A⟦ρ,σ,a⟧ op □⟩∷κl']
⟨a,ρ,σ,κl,κσ,τ⟩ --> ⟨e,ρ'',σ',κl',κσ,τ+1⟩
  where ⟨⟨λ(x).e,ρ'⟩ @ □⟩∷κl':= κσ(κl)
        σ' := σ[(x,τ) ↦ A⟦ρ,σ,a⟧]
        ρ'' := ρ'[x ↦ (x,τ)]
⟨i₂,ρ,σ,κl,κσ,τ⟩ --> ⟨i,ρ,σ,κl',κσ,τ+1⟩
  where ⟨i₁ iop □⟩∷κl' := κσ(κl)
        i := δ⟦iop,i₁,i₂⟧
⟨i,ρ,σ,κl,κσ,τ⟩ --> ⟨e,ρ,σ,κl',κσ,τ+1⟩
  where ⟨if0(□){e₁}{e₂}⟩∷κl' := κσ(κl)
        e := if(i = 0) then e₁ else e₂
```

We also wish to employ abstract garbage collection, which adheres to the following specification:

```
_~~>_ ∈ P(Σ × Σ)
ς ~~> ς'
  where ς --> ς'
⟨e,ρ,σ,κl,κσ,τ⟩ ~~> ⟨e,ρ,σ',κl,κσ,τ⟩
  where σ' := {l ↦ σ(l) | l ∈ R[σ](ρ,e)}
        κσ' := {κl ↦ κσ(κl) | κl ∈ κR[κσ](κl)}
```

where `R` is the set of addresses reachable from a given expression:

```
R[_] ∈ Store → Env × Exp → P(Addr)
R[σ](ρ,e) := μ(θ). R₀(ρ,e) ∪ θ ∪ {l' | l' ∈ R-Val(σ(l)) ; l ∈ θ}

R₀ ∈ Env × Exp → P(Addr)
R₀(ρ,e) := {ρ(x) | x ∈ FV(e)}

FV ∈ Exp → P(Var)
FV(x) := {x}
FV(i) := {}
FV(λ(x).e) := FV(e) - {x}
FV(e₁ op e₂) := FV(e₁) ∪ FV(e₂)
FV(if0(e₁){e₂}{e₃}) := FV(e₁) ∪ FV(e₂) ∪ FV(e₃)

R-Val ∈ Val → P(Addr)
R-Val(i) := {}
R-Val(⟨λ(x).e,ρ⟩) := {ρ(x) | y ∈ FV(λ(x).e)}
```

`R[σ](ρ,e)` computes the transitively reachable addresses from `e` in `ρ` and `σ`. (We write `μ(x). f(x)` as the least-fixed-point of a function `f`.) `R₀(ρ,e)` computes the initial reachable address set for `e` under `ρ`. `FV(e)` computes the free variables for an expression `e`. `R-Val` computes the addresses reachable from a value.

Analagously, `κR` is the set of addresses reachable from a given continuation address:

```
κR[_] ∈ KStore → KAddr → P(KAddr)
κR[κσ](κl) := μ(kθ). κθ₀ ∪ κθ ∪ { π₂(κσ(κl)) | κl ∈ κθ}
```

## 3. Monadic Interpreter

We next design an interpreter for `λIF` as a monadic interpreter. This interpreter will support both concrete and abstract executions. To do this, there will be three parameters which the user can instantiate in any way they wish:

1. The monad, which captures the flow-sensitivity of the analysis.
2. The value space, which captures the abstract domain for integers and closures.
3. Abstract time, which captures the call-site sensitivity of the analysis.

We place each of these features behind an abstract interface and leave theirl implementations opaque. We will recover specific concrete and abstract interpreters in a later section.

The goal is to implement as much of the interpreter as possible while leaving these things abstract. The more we can prove about the interpreter independent of these variables, the more proof-work we'll get for free.

### 3.1 The Monad Interface

The interpreter will use a monad `M` in two ways. First, to manipulate components of the state space (like `Env` and 'Store). Second, to exhibit nondeterministic behavior, which is inherent in computable analysis. We capture these properties as monadic effects.

To be a monad, `M` must have type:

```
M : Type → Type
```

and support the `bind` operation:

```
bind : ∀ α β, M(α) → (α → M(β)) → M(β)
```

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

```
return : ∀ α, α → M(α)
```

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

```
bind-unit₁ : bind(return(a))(k) = k(a)
bind-unit₂ : bind(m)(return) = m
bind-associativity :
  bind(bind(m)(k₁))(k₂) = bind(m)(λ(a)→bind(k₁(a))(k₂))
```

These operators capture the essence of the explicit state-passing and set comprehension aspects of the interpreter. Our interpreter will use these operators and avoid referencing an explicit configuration ς or sets of results.

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

```
do
  a ← m
  k(a)
```

and

```
a ← m ; k(a)
```

are both just sugar for

```
bind(m)(k)
```

Interacting with `Env` is achieved through `get-Env` and `put-Env` effects:

```
get-Env : M(Env)
put-Env : Env → M(1)
```

which have the following laws:

```
put-put : put-Env(s₁) ; put-Env(s₂) = put-Env(s₂)
put-get : put-Env(s) ; get-Env = return(s)
get-put : s ← get-Env ; put-Env(s) = return(1)
get-get : s₁ ← get-Env ; s₂ ← get-Env ; k(s₁,s₂) = s ← get-Env ; k(s,s)
```

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

Nondeterminism is achieved through operators `⟨⊥⟩` and `⟨+⟩`:

```
⟨⊥⟩ : ∀ α, M(α)
_⟨+⟩_ : ∀ α, M(α) × M(α) → M α
```

which have the following laws:

```
⊥-zero₁ : bind(⟨⊥⟩)(k) = ⟨⊥⟩
⊥-zero₂ : bind(m)(λ(a)→⟨⊥⟩) = ⟨⊥⟩
⊥-unit₁ : ⟨⊥⟩ ⟨+⟩ m = m
⊥-unit₂ : m ⟨+⟩ ⟨⊥⟩ = m
+-associativity : m₁ ⟨+⟩ (m₂ ⟨+⟩ m₃) = (m₁ ⟨+⟩ m₂) ⟨+⟩ m₃
+-commutativity : m₁ ⟨+⟩ m₂ = m₂ ⟨+⟩ m₁
+-distributivity : bind(m₁ ⟨+⟩ m₂)(k) = bind(m₁)(k) ⟨+⟩ bind(m₂)(k)
```

### 3.2 The Value Space Interface

To abstract the value space we require the type `Val` be an opaque parameter We need only require that `Val` is a join-semilattice:

```
⊥ : Val
_`join`_ : Val × Val → Val
```

The interface for integers consists of introduction and elimiation rules:

```
int-I : ℤ → Val
int-if0-E : Val → P(Bool)
```

We can now state laws for this interface, which are designed to induce a Galois connection between `ℤ` and `Val`:

```
{true}  ⊑ int-if0-E(int-I(i))     if i = 0
{false} ⊑ int-if0-E(int-I(i))     if i ≠ 0


v ⊒ `bigjoin`_{b ∈ int-if0-E(v)} θ(b)
  where θ(true)  = int-I(0)
        θ(false) = `bigjoin`_{i ∈ ℤ | i ≠ 0} int-I(i)
```

Additionally we must abstract closures:

```
clo-I : Clo → Val
clo-E : Val → P(Clo)
```

which follow similar laws:

```
{c} ⊑ clo-E(cloI(c))
v ⊑ `bigjoin`_{c ∈ clo-E(v)} clo-I(c)
```

The denotation for primitive operations `δ` must also be opaque:

```
δ : IOp × Val × Val → Val
```

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.

### 3.3 Abstract Time

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

```
tick : Exp × KAddr × Time → 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 `Exp × KAddr`. Also in AAM is the opaque function `alloc : Var × Time → 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.

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`:

```
σ  ∈ Store  : Addr → Val
κσ ∈ KStore : KAddr → P(Frame × 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 `P(Frame × KAddr)` as an abstraction for continuations for simplicity.

### 3.4 Interpreter Definition

We use the three interfaces from above as opaque parameters to out interpreter. Before defining the interpreter we define three helper functions. These helper functions crucially rely on the monadic effect interface.

First, values in `P(α)` can be lifted to monadic values `M(α)` using `return` and `⟨⊥⟩`, which we name `↑ₚ`:

```
↑ₚ : ∀ α, P(α) → M(α)
↑ₚ({a₁ .. aₙ}) := return(a₁) ⟨+⟩ .. ⟨+⟩ return(aₙ)
```

We introduce monadic helper functions for allocation and manipulating time:

```
allocM : Var → M(Addr)
allocM(x) := do
  τ ← get-Time
  return(x,τ)

κallocM : M(KAddr)
κallocM := do
  τ ← get-Time
  return(τ)


tickM : Exp → M(1)
tickM(e) = do
  τ ← get-Time
  κl ← get-KAddr
  put-Time(tick(e,κl,τ))
```

Finally we introduce helper functions for manipulating stack frames:

```
push : Frame → M(1)
push(fr) := do
  κl ← get-KAddr
  κσ ← get-KStore
  κl' ← κallocM
  put-KStore(κσ `join` [κl' ↦ {fr∷κl}])
  put-KAddr(κl')

pop : M(Frame)
pop := do
  κl ← get-KAddr
  κσ ← get-KStore
  fr∷κl' ← ↑ₚ(κσ(κl))
  put-KAddr(κl')
  return(fr)
```

We can now write a monadic interpreter for `λIF` using these monadic effects.

```
A⟦_⟧ ∈ Atom → M(Val)
A⟦i⟧ := return(int-I(i))
A⟦x⟧ := do
  ρ ← get-Env
  σ ← get-Store
  l ← ↑ₚ(ρ(x))
  return(σ(x))
A⟦λ(x).e⟧ := do
  ρ ← get-Env
  return(clo-I(⟨λ(x).e,ρ⟩))

step : Exp → M(Exp)
step(e₁ op e₂) := do
  tickM(e₁ op e₂)
  push(⟨□ op e₂⟩)
  return(e₁)
step(a) := do
  tickM(a)
  f ← pop
  v ← A⟦a⟧
  case f of
    ⟨□ op e⟩ → do
      push(⟨v op □⟩)
      return(e)
    ⟨v' @ □⟩ → do
      ⟨λ(x).e,ρ'⟩ ← ↑ₚ(clo-E(v'))
      l ← alloc(x)
      σ ← get-Store
      put-Env(ρ'[x↦l])
      put-Store(σ[l↦v])
      return(e)
    ⟨v' iop □⟩ → do
      return(δ(iop,v',v))
    ⟨if0(□){e₁}{e₂}⟩ → do
      b ← ↑ₚ(int-if0-E(v))
      if(b) then return(e₁) else return(e₂)
```

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 `(Σ → Σ)α⇄γ(Exp → M(Exp))`. This Galois connection serves two purposes. First, it allows us to implement the analysis by converting our interpreter to the transition system `Σ → Σ` through `γ`. Second, this Galois connection serves to *transport other Galois connections*. For example, given concrete and abstract versions of `Val`, we carry `CVal α⇄γ AVal` through the Galois connection to establish `CΣ α⇄γ AΣ`.

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

```
μ(ς). ς₀ `join` ς `join` γ(step)(ς)
```

where `ς₀` is the injection of the initial program `e` into `Σ`.

## 4. Recovering Concrete and Abstract Interpreters

To recover a concrete interpreter we instantiate `M` to a path-sensitive monad: `M`$^{ps}$. The path sensitive monad is a simple powerset of products:

```
ψ ∈ Ψᵖˢ := Env × Store × KAddr × KStore × Time
m ∈ Mᵖˢ(α) := Ψᵖˢ → P(α × Ψᵖˢ)
```

Monadic operators `bind`$^{ps}$ and `return`$^{ps}$ are defined to encapsulate both state-passing and set-flattening:

```
bindᵖˢ : ∀ α, Mᵖˢ(α) → (α → Mᵖˢ(β)) → Mᵖˢ(β)
bindᵖˢ(m)(k)(ψ) := {(y,ψ'') | (y,ψ'') ∈ k(a)(ψ') ; (a,ψ') ∈ m(ψ)}

returnᵖˢ : ∀ α, α → Mᵖˢ(α)
returnᵖˢ(a)(ψ) := {(a,ψ)}
```

State effects merely return singleton sets:

```
get-Envᵖˢ : Mᵖˢ(Env)
get-Envᵖˢ(⟨ρ,σ,κ,τ⟩) := {(ρ,⟨ρ,σ,κ,τ⟩)}

put-Envᵖˢ : Env → P(1)
put-Envᵖˢ(ρ')(⟨ρ,σ,κ,τ⟩) := {(1,⟨ρ',σ,κ,τ⟩)}
```

Nondeterminism effects are implemented with set union:

```
⟨⊥⟩ᵖˢ : ∀ α, Mᵖˢ(α)
⟨⊥⟩ᵖˢ(ψ) := {}

_⟨+⟩ᵖˢ_ : ∀ α, Mᵖˢ(α) × Mᵖˢ(α) → Mᵖˢ(α)
(m₁ ⟨+⟩ᵖˢ m₂)(ψ) := m₁(ψ) ∪ m₂(ψ)
```

*Proposition: M satisfies monad, state, and nondeterminism laws.*

For the value space `CVal` we use a powerset of semantic values `Val`:

```
v ∈ CVal := P(Val)
```

with introduction and elimination rules:

```
int-I : ℤ → CVal
int-I(i) := {i}

int-if0-E : CVal → P(Bool)
int-if0-E(v) := { true | 0 ∈ v } ∪ { false | i ∈ v ∧ i ≠ 0 }
```

and `δ` to manipulate abstract values:

```
δ⟦_,_,_⟧ : IOp × CVal × CVal → CVal
δ⟦+,v₁,v₂⟧ := { i₁ + i₂ | i₁ ∈ v₁ ; i₂ ∈ v₂ }
δ⟦-,v₁,v₂⟧ := { i₁ - i₂ | i₁ ∈ v₁ ; i₂ ∈ v₂ }
```

Abstract time and addresses are program contours in the concrete space:

```
τ  ∈ Time  := (Exp × KAddr)*
l  ∈ Addr  := Var × Time
κl ∈ KAddr := Time
```

Operators `alloc` and `κalloc` are merely identity functions, and `tick` is just a cons operator.

Finally, we must establish a Galois connection between `Exp → Mᵖˢ(Exp)` and `Σ → Σ` for some `Σ`. The state space `Σ` depends only on the monad `M`$^{ps}$ and is independent of the choice for `CVal`, Addr or Time. For the path sensitive monad `M`$^{ps}$ , `Σ`$^{ps}$ is defined:

```
Σᵖˢ := P(Exp × Ψᵖˢ)
```

and the Galois connection is:

```
γᵖˢ : (Exp → Mᵖˢ(Exp)) → Σᵖˢ → Σᵖˢ
γᵖˢ(f)(eψ*) := {(e,ψ') | (e,ψ') ∈ f(e)(ψ) ; (e,ψ) ∈ eψ*}


αᵖˢ : (Σᵖˢ → Σᵖˢ) → Exp → Mᵖˢ(Exp)
αᵖˢ(f)(e)(ψ) := f({(e,ψ)})
```

*Proposition: $\gamma^{ps}$ and $\alpha^{ps}$ form an isomorphism.* This implies Galois connnection.

The injection $\varsigma^{ps}_0$ for a program `e` is:

```
ςᵖˢ₀ := {⟨e,⊥,⊥,•,⊥,•⟩}
```

To arrive at an abstract interpreter we seek a finite state space. First we abstract the value space `Val` as `AVal`, which only tracks integer parity:

```
AVal := P(Clo + {-,0,+})
```

Introduction and elimination functions are defined:

```
int-I : ℤ → AVal
int-I(i) := - if i < 0
            0 if i = 0
            + if i > 0


int-if0-E : AVal → P(Bool)
int-if0-E(v) := { true | 0 ∈ v } ∪ { false | - ∈ v ∨ + ∈ v }
```

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

The abstract `δ` operator is defined:

```
Aδ : IOp × AVal × AVal → AVal
Aδ(+,v₁,v₂) := { p      | 0 ∈ v₁ ∧ p ∈ v₂ }
             ∪ { p      | p ∈ v₁ ∧ 0 ∈ v₂ }
             ∪ { +      | + ∈ v₁ ∧ + ∈ v₂ }
             ∪ { -      | - ∈ v₁ ∧ - ∈ v₂ }
             ∪ { -,0,+ | + ∈ v₁ ∧ - ∈ v₂ }
             ∪ { -,0,+ | - ∈ v₁ ∧ + ∈ v₂ }
```

Next we abstract `Time` to the finite domain of a k-truncated list of execution contexts:

```
Time := (Exp × KAddr)*ₖ
```

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

```
tick : Exp × KAddr × Time → Time
tick(e,κl,τ) = ⌊(e,κl)∷τ⌋ₖ
```

After substituting abstract versions for `Val` and `Time`, the following state space for `Σᵖˢ` becomes finite:

```
P(Exp × AEnv × AStore × AKAddr × AKStore × ATime)
```

and the least-fixed-point iteration of the collecting semantics provides a sound and computable analysis.

## 5. Varying Path and Flow Sensitivity

We are able to recover a flow-insensitive interpreter through a new definition for `M`: `Mᶠⁱ`. To do this we pull `Store` out of the powerset and use its join-semilattice structure:

```
Ψᶠⁱ := Env × KAddr × KStore × Time
Mᶠⁱ(α) := Ψᶠⁱ × Store × P(α × Ψᶠⁱ) × Store
```

The monad operator `bindᶠⁱ` must merge multiple stores back to one:

```
bindᶠⁱ : ∀ α β, Mᶠⁱ(α) → (α → Mᶠⁱ(β)) → Mᶠⁱ(β)
bindᶠⁱ(m)(k)(ψ,σ) := ({bs₁₁ .. bsₙ₁ .. bsₙₘ},σ₁ `join` .. `join` σₙ)
  where
    ({(a₁,ψ₁) .. (aₙ,ψₙ)},σ') := m(ψ,σ)
    ({bψᵢ₁ .. bψᵢₘ},σᵢ) := k(aᵢ)(ψᵢ,σ')
```

The unit for `bindᶠⁱ`:

```
returnᶠⁱ : ∀ α, α → Mᶠⁱ(α)
returnᶠⁱ(a)(ψ,σ) := ({a,ψ},σ)
```

State effects `get-Env` and `put-Env`:

```
get-Envᶠⁱ : Mᶠⁱ(Env)
get-Envᶠⁱ(⟨ρ,κ,τ⟩,σ) := ({(ρ,⟨ρ,κ,τ⟩)},σ)


put-Envᶠⁱ : Env → Mᶠⁱ(1)
put-Envᶠⁱ(ρ')(⟨ρ,κ,τ⟩,σ) := ({(1,⟨ρ',κ,τ⟩)},σ)
```

State effects `get-Store` and `put-Store`:

```
get-Storeᶠⁱ : Mᶠⁱ(Env)
get-Storeᶠⁱ(⟨ρ,κ,τ⟩,σ) := ({(σ,⟨ρ,κ,τ⟩},σ)


put-Storeᶠⁱ : Store → Mᶠⁱ(1)
put-Storeᶠⁱ(σ')(⟨ρ,κ,τ⟩,σ) := ({(1,⟨ρ,κ,τ⟩)},σ')
```

Nondeterminism operations:

```
⟨⊥⟩ᶠⁱ : ∀ α, M(α)
⟨⊥⟩ᶠⁱ(ψ,σ) := ({}, ⊥)


_⟨+⟩_ : ∀ α, M(α) × M(α) → M α
(m₁ ⟨+⟩ m₂)(ψ,σ) := (aψ*₁ ∪ aψ*₂,σ₁ `join` σ₂)
  where (aψ*ᵢ,σᵢ) := mᵢ(ψ,σ)
```

Finally, the Galois connection for relating `Mᶠⁱ` to a state space transition over `Σᶠⁱ`:

```
Σᶠⁱ := P(Exp × Ψᶠⁱ) × Store


γᶠⁱ : (Exp → Mᶠⁱ(Exp)) → (Σᶠⁱ → Σᶠⁱ)
γᶠⁱ(f)(eψ*,σ) := ({eψ₁₁ .. eψₙ₁  .. eψₙₘ}, σ₁ `join` .. `join` σₙ)
  where {(e₁,ψ₁) .. (eₙ,ψₙ)} := eψ*
        ({eψᵢ₁ .. eψᵢₘ},σᵢ) := f(eᵢ)(ψᵢ,σ)


αᶠⁱ  : (Σᶠⁱ → Σᶠⁱ) → (Exp → Mᶠⁱ(Exp))
αᶠⁱ(f)(e)(ψ,σ) := f({(e,ψ)},σ)
```

*Proposition:* $\gamma^{fi}$ *and* $\alpha^{fi}$ *form an isomorphism.* Like the concrete $\gamma^{fi}$ and $\alpha^{fi}$, this implies Galois connection.

*Proposition:* $M^{ps}$ $\alpha\rightleftarrows\gamma$ $M^{fi}$*.* This demonstrates that path sensitivity is more precise than flow insensitivity in a formal, language-independent setting.

We leave out the explicit definition for the flow-sensitive monad $M^{fs}$. However, we will recover it through the compositional framework in Section [X][A Compositional Framework] using monad transformers.

}}}

## 6. A Compositional Monadic Framework

In our framework thus far, any modification to the interpreter requires redesigning the monad M. However, we want to avoid reconstructing complicated monads for our interpreters. Even more, we want to avoid reconstructing *proofs* about monads for our interpreters. Toward this goal we introduce a compositional framework for constructing monads 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:

```
Sₜ[_] : (Type → Type) → (Type → Type)
Sₜ[s](m)(α) := s → m(α × s)
```

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

```
bindˢ : ∀ α β, Sₜ[s](m)(α) → (α → Sₜ[s](m)(β)) → Sₜ[s](m)(β)
bindˢ(m)(k)(s) := do
  (x,s') ←ᵐ m(s)
  k(x)(s')

returnˢ : ∀ α m, α → Sₜ[s](m)(α)
returnˢ(x)(s) := returnᵐ(x,s)
```

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

```
getˢ : Sₜ[s](m)(s)
getˢ(s) := returnᵐ(s,s)

putˢ : s → Sₜ[s](m)(1)
putˢ(s')(s) := returnᵐ(1,s')
```

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

```
⟨⊥⟩ : ∀ α, Sₜ[s](m)(α)
⟨⊥⟩(s) := ⟨⊥⟩ᵐ

_⟨+⟩_ : ∀ α, Sₜ[s](m)(α) x Sₜ[s](m)(α) → Sₜ[s](m)(α)
(m₁ ⟨+⟩ m₂)(s) := m₁(s) ⟨+⟩ᵐ m₂(s)
```

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.

```
Pₜ : (Type → Type) → (Type → Type)
Pₜ(m)(α) := m(P(α))
```

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

```
bindᵖ : ∀ α β, Pₜ(m)(α) → (α → Pₜ(m)(β)) → Pₜ(m)(β)
bindᵖ(m)(k) := do
  {x₁ .. xₙ} ←ᵐ m
  k(x₁) `join`ᵐ .. `join`ᵐ k(xₙ)

returnᵖ : ∀ α, α → Pₜ(m)(α)
returnᵖ(x) := returnᵐ({x})
```

Nondterminism actions $\langle\bot\rangle^m$ and $+$ interact with the join-semilattice functorality of the underlying monad m:

```
⟨⊥⟩ᵖ : ∀ α, Pₜ(m)(α)
⟨⊥⟩ᵖ := ⊥ᵐ

_⟨+⟩_ : ∀ α, Pₜ(m)(α) x Pₜ(m)(α) → Pₜ(m)(α)
m₁ ⟨+⟩ᵖ m₂ := m₁ `join`ᵐ m₂
```

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

```
getᵖ : Pₜ(m)(s)
getᵖ = mapᵖ(λ(s).{s})(getᵐ)

putᵖ : s → Pₜ(m)(s)
putᵖ(s) = mapᵖ(λ(1).{1})(putᵐ(s))
```

*Proposition:* $P_t$ *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 : ∀ α β, (α → M(β)) α⇄γ (Σ(α) → Σ(β))
```

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

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

```
mstepˢ-γ : ∀ α β m, (α → Sₜ[s](m)(β)) → (Σᵐ(α × s) → Σᵐ(β × s))
mstepˢ-γ(f) := mstepᵐ-γ(λ(a,s). f(a)(s))
```

For the nondeterminism transformer $P_t$, mstep has two possible definitions. One where Σ is $\Sigma^m \circ P$:

```
mstepᵖ₁-γ : ∀ α β m, (α → Pₜ(m)(β)) → (Σᵐ(P(α)) → Σᵐ(P(β)))
mstepᵖ₁-γ(f) := mstepᵐ-γ(λ({x₁ .. xₙ}). f(x₁) ⟨+⟩ .. ⟨+⟩ f(xₙ))
```

and one where $\Sigma$ is $P \circ \Sigma^m$:

```
mstepᵖ₂-γ : ∀ α β m, (α → Pₜ(m)(β)) → (P(Σₘ(α)) → P(Σₘ(β)))
mstepᵖ₂-γ(f)({ς₁ .. ςₙ}) := aΣP₁ ∪ .. ∪ aΣPₙ
  where
    commuteP : ∀ α, Σᵐ(P(α)) → P(Σᵐ(α))
    aΣPᵢ := commuteP-γ(mstepᵐ-γ(f)(ςᵢ))
```

The operation `computeP` must be defined for the underlying $\Sigma^m$. This property is true for the identiy monad, and is preserved by $S_t[s]$ when $\Sigma^m$ is also a functor:

```
commuteP-γ : ∀ α, Σᵐ(P(α) × s) → P(Σᵐ(α × s))
commuteP-γ := commutePᵐ ∘ map(λ({α₁ .. αₙ},s). {(α₁,s) .. (αₙ,s)})
```

The $\gamma$ side of commuteP is the only Galois connection mapping that loses information in the $\alpha$ direction. Therefore, $\mathtt{mstep}^{s}$ and $\mathtt{mstep}^{p}_{1}$ are really isomorphism transformers, and $\mathtt{mstep}^{p}_{2}$ is the only Galois connection transformer.

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

For convenience, we name the pairing of $P_t$ with $\mathtt{mstep}^{p}_{1}$ $FI_t$, and with $\mathtt{mstep}^{p}_{2}$ $FS_t$ for flow insensitive and flow sensitive respectively.

We can now build monad transformer stacks from combinations of $S_t[s]$, $FI_t$ and $FS_t$ 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 $\Sigma$.
- Galois connections between states $s_1$ and $s_2$ are transported along the Galois connection between $(\alpha \to S_t[s_1](m)(\beta)) \; \alpha{\rightleftarrows}\gamma \; (\Sigma[s_1](\alpha) \to \Sigma[s_1](\beta))$ and $(\alpha \to S_t[s_2](m)(\beta)) \; \alpha{\rightleftarrows}\gamma \; (\Sigma[s_2](\alpha) \to \Sigma[s_2](\beta))$ resulting in $(\Sigma[s_1](\alpha) \to \Sigma[s_1](\beta)) \; \alpha{\rightleftarrows}\beta \; (\Sigma[s_2](\alpha) \to \Sigma[s_2](\beta))$.

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

- $S_t[Env] \circ S_t[Store] \circ S_t[KAddr] \circ S_t[KStore] \circ S_t[Time] \circ FS_t$
  - This yields a path-sensitive flow-sensitive analysis.
- $S_t[Env] \circ S_t[KAddr] \circ S_t[KStore] \circ S_t[Time] \circ FS_t \circ S_t[Store]$
  - This yeilds a path-insensitive flow-sensitive analysis.
- $S_t[Env] \circ S_t[KAddr] \circ S_t[KStore] \circ S_t[Time] \circ FI_t \circ S_t[Store]$
  - This yields a path-insensitive flow-insensitive analysis.

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