# Hiding local state in direct style: a higher-order anti-frame rule

François Pottier

January 28th, 2008



#### Contents

- Introduction
- Basics of the type system
- A higher-order anti-frame rule
- Applications
- Conclusion
- Bibliography

### Hidden state

Many "objects" (or "modules", "components", "functions") rely on a piece of modifiable internal state, yet publish an informal specification that does not reveal the existence of such a state.

### Hidden state

For instance, a memory manager might maintain a linked list of freed memory blocks.

Yet, clients need not know anything about it.

It is safe for them to consider that the allocation and deallocation functions have no side effect, other than the obvious effect of providing the client with, or depriving the client from, ownership of a unique memory block.

Hiding is not abstraction. Hiding pretends that there is no internal state, while abstraction acknowledges that there is one, but makes its type (and properties) abstract.

Both protect the internal state from interference by clients, and protect clients from changes in the representation of the internal state.

Hiding offers the additional advantage that objects with internal state appear as ordinary objects, hence can be untracked. It is not necessary to ask how they are aliased or who owns them.

Abstraction offers the additional advantage that clients can reason about state changes. The computational state, which has abstract type, can be declared to represent some logical state, at a concrete type. For instance, the internal state of a hash table represents a mathematical finite map.

In practice, both hiding and abstraction are useful.

Consider an object that produces the stream of the prime numbers.

If it is specified that each invocation returns the *next* prime number, then the internal state can only be *abstract*.

If it is only specified that each invocation returns some prime number, then the state can be hidden.

Whether an object's internal state can be hidden depends not just on the object's actual behavior, but also on its specification.

As specifications become less precise, opportunities for hiding state increase!

### Towards a formalization

How could the concept of hidden state be made precise in a formal framework for reasoning about programs?

This talk attempts to provide an answer...

### Towards a formalization

Which formal frameworks provide an appropriate setting in which to ask (and answer) this question?

There are several. Separation logic is one. A type system with regions and capabilities is another.

In fact, the two are quite close: both keep track of aliasing and ownership properties. Both allow assigning pre- and post-conditions to code.

### Towards a formalization

Hidden state has been previously studied in the setting of separation logic [O'Hearn et al., 2004, Birkedal et al., 2006].

In this talk, I use the vocabulary of a type system [Charguéraud and Pottier, 2007] for an ML-like programming language.

It should be possible to transpose the main idea to the setting of separation logic.

#### Contents

- Introduction
- Basics of the type system
- A higher-order anti-frame rule
- Applications
- Conclusion
- Bibliography

## The host type system

This type system is the setting in which I develop a rule for hiding state and prove (syntactic) type soundness.

The details of the type system are somewhat unimportant for this talk. I just wish to convey a flavor of the system before we embark on a journey towards hidden state...

## Regions

A region  $\rho$  is a static name for a set of values.

The type [
ho] is the type of the values that inhabit the region ho.

In this talk, there are only singleton regions, so a region  $\rho$  is a static name for a value, and  $[\rho]$  is a singleton type. My paper with Arthur Charguéraud [2007] also has group regions, which become necessary when there is aliasing.

## Capabilities

A singleton capability  $\{\rho:\theta\}$  is a static token that serves two roles.

First, it carries a memory type  $\theta$ , which describes the structure and extent of the memory area to which the value  $\rho$  gives access. Second, it represents ownership of this area.

For instance,  $\{\rho: \text{ref int}\}$  asserts that the value  $\rho$  is the address of a reference cell, and asserts ownership of this cell.

## Capabilities (summary)

On top of singleton capabilities, one builds composite capabilities:

$$\begin{array}{lll} \textit{C} & ::= & \emptyset & & \text{empty heap} \\ & \mid & \{\rho:\theta\} & & \text{singleton heap} \\ & \mid & C_1 \land C_2 & & \text{(separating) conjunction} \\ & \mid & \exists \rho. \textit{C} & & \text{embedded region} \\ & \mid & C_1 \otimes C_2 & & \text{(explained later on)} \end{array}$$

There is a clear analogy between capabilities and separation logic assertions.

Here is a summary of memory types:

$$\begin{array}{lll} \theta & ::= & \bot \mid \text{unit} \mid \theta_1 + \theta_2 \mid \theta_1 \times \theta_2 & \text{data} \\ & \mid & \sigma_1 \to \sigma_2 & \text{functions} \\ & \mid & [\rho] & \text{indirection via a region} \\ & \mid & \text{ref } \theta & \text{reference cell} \\ & \mid & C \wedge \theta & \text{embedded capability} \\ & \mid & \exists \rho.\theta & \text{embedded region} \\ & \mid & \theta \otimes C & \text{(explained later on)} \end{array}$$

Memory types express ownership, so they are linear.

## Type-level recursion

I assume that capabilities and types can be recursive. For instance, the unique solution to the following equation:

$$R = \{ \rho : ref((R \wedge \sigma_1) \rightarrow (R \wedge \sigma_2)) \}$$

is a singleton capability R for a reference cell that contains a function that requires, and preserves, R.

Recursive capabilities are required in many applications, and, in this system, appear necessary for the subject reduction proof to go through.

## Value types (summary)

#### Values receive value types:

Values are non-linear: they can be discarded or duplicated at will.

Value types form a subset of memory types, deprived of references and embedded capabilities.

## Judgements about values

Judgements about values take the form:

$$\Gamma \vdash V : \tau$$

Type environments  $\Gamma$  associate value types with variables.

Values do not involve computation, which is why this judgement form does not involve any capabilities, either as input or as output.

## Judgements about terms

Judgements about terms take the form:

$$\Gamma$$
;  $C \vdash t : \sigma$ 

The capability C serves a pre-condition, while the computation type  $\sigma$  serves as a post-condition. Judgements about terms are analogous to Hoare triples in separation logic.

Computation types are:

$$\sigma ::= \tau \mid C \land \sigma \mid \exists \rho . \sigma \mid \sigma \otimes C$$

# Typing rules for references

References are tracked: allocation produces a singleton capability, which is later required for access. Read and write accesses are restricted to non-linear value types, because they duplicate or discard a value.

```
ref : \tau \to \exists \rho. \{\rho : \text{ref } \tau\} [\rho]

get : \{\rho : \text{ref } \tau\} [\rho] \to \{\rho : \text{ref } \tau\} \tau

set : \{\rho : \text{ref } \tau_1\} ([\rho] \times \tau_2) \to \{\rho : \text{ref } \tau_2\} \text{ unit}
```

References to non-value types can be created and exploited via focusing:

(focus-ref) 
$$\{\rho_1 : \text{ref } \theta\} \equiv \exists \rho_2. \{\rho_1 : \text{ref } [\rho_2]\} \{\rho_2 : \theta\}$$

#### Contents

- Introduction
- Basics of the type system
- A higher-order anti-frame rule
- Applications
- Conclusion
- Bibliography

#### The first-order frame rule

The first-order *frame rule* states that, if a term behaves correctly in a certain store, then it also behaves correctly in a larger store, and does not affect the part of the store that it does not know about:

$$\frac{\Gamma; C_2 \vdash t : \sigma}{\Gamma; (C_1 \land C_2) \vdash t : (C_1 \land \sigma)}$$

This rule can also take the form of a simple subtyping axiom:

$$\sigma_1 \rightarrow \sigma_2 \leq (C \wedge \sigma_1) \rightarrow (C \wedge \sigma_2)$$

### The first-order frame rule

The frame rule makes a capability unknown to a term, while known to its context.

To hide a piece of local state is the exact dual: to make a capability known to a term, yet unknown to its context.

In a programming language with higher-order functions, one could hope to be able to exploit the duality between terms and contexts.

By viewing the context as a term, a continuation, one could perhaps use a frame rule to hide a piece of local state.

This is the approach of Birkedal, Torp-Smith, and Yang [2006], who follow up on earlier work by O'Hearn, Yang, and Reynolds [2004].

Imagine that we have a provider, a term of type:

$$C \wedge ((C \wedge unit) \rightarrow (C \wedge int))$$

The provider initially establishes  ${\it C}$  and returns a function that requires  ${\it C}$  and preserves it.

This could be the type of a stream of integers, with internal state.

We now wish to hide C and pretend that the provider is an ordinary function, of type unit  $\rightarrow$  int.

Applying the frame rule to the provider would not help.

We must apply the frame rule to the client, assuming it is known.

Imagine that we also have a client, a term of type:

$$(unit \rightarrow int) \rightarrow a$$

This client is explicitly abstracted over the provider. The type a is some answer type.

The client does not know about the invariant  $\mathcal{C}$ . It views the provider as an ordinary function, without side effects.

At first, the function application (client provider) seems ill-typed. The provider offers:

$$(C \land unit) \rightarrow (C \land int)$$

while the client requires:

unit 
$$\rightarrow$$
 int

The former is not a subtype of the latter: in fact, according to the first-order frame rule, it is the other way around!

This is where Birkedal et al.'s higher-order frame rule [2006] comes into play. The rule guarantees:

$$(\mathsf{unit} \to \mathsf{int}) \to a \quad \leq \quad (C \land (C \land \mathsf{unit} \to C \land \mathsf{int})) \to (C \land a)$$

That is, if C holds initially and if the provider preserves C, then, the client will unwittingly preserve it as well.

Here, the first-order frame rule would yield a weaker statement:

$$(unit \rightarrow int) \rightarrow a \leq (C \land (unit \rightarrow int)) \rightarrow (C \land a)$$

# The higher-order frame rule

The general form of the higher-order frame rule is:

$$\sigma \leq \sigma \otimes C$$

The type  $\sigma \otimes C$  (" $\sigma$  under C") describes the same behavior as  $\sigma$ , and additionally requires C to be available at every interaction between the term and its context.

# The higher-order frame rule

The operator  $\cdot \otimes C$  makes C a new pre-condition and a new post-condition of every arrow within its left-hand argument:

$$(\sigma_1 \to \sigma_2) \otimes C = (C \wedge (\sigma_1 \otimes C)) \to (C \wedge (\sigma_2 \otimes C))$$

The operator  $\cdot \otimes C$  commutes with products, sums, references, etc. It vanishes at base types.

After applying the higher-order frame rule, the client has type:

$$(C \land (C \land unit \rightarrow C \land int)) \rightarrow (C \land a)$$

Recall that the provider has type:

$$C \wedge ((C \wedge unit) \rightarrow (C \wedge int))$$

So the function application (client provider) is in fact well-typed, and has type  $C \wedge a$ .

In a modular setting, the client is unknown. One can then abstract the provider over the client. If one admits the subtyping axiom  $C \leq \emptyset$ , then the value:

has type:

$$((unit \rightarrow int) \rightarrow a) \rightarrow a$$

This is the double negation of the desired type.

We succeeded, but were led to use continuation-passing style.

Is this approach to hidden state realistic?

I claim not: continuation-passing style is not practical.

What is a direct-style analogue of the higher-order frame rule?

## Towards a higher-order anti-frame rule

We need a higher-order anti-frame rule, that is, a rule that applies not to the term, but to its context, without requiring an explicit switch to continuation-passing style.

# Towards a higher-order anti-frame rule

An approximation of such a rule is:

$$C \wedge (\sigma \otimes C) \leq \sigma$$
 (unsound)

The left-hand side of the rule states that:

- Term must guarantee C when abandoning control to Context;
- Term may assume C when receiving control from Context;

In that case, it should be safe for Context to not know about C. The intended invariant is, C holds whenever Context has control.

# Towards a higher-order anti-frame rule

The candidate rule on the previous slide is sound only for *closed* terms that run in an *empty* store.

In general, interaction between Term and Context takes place also via the (function) values that can be reached via the environment or the store.

As a result, the type environment and the type of the store too must come in two versions. Term's view is that  $\mathcal C$  holds at every interaction, while Context's view does not even mention  $\mathcal C$ .

# A higher-order anti-frame rule

A sound version of the rule is:

Anti-frame
$$\frac{\Gamma \otimes C_1; C_2 \otimes C_1 \vdash t : C_1 \land (\sigma \otimes C_1)}{\Gamma; C_2 \vdash t : \sigma}$$

This is dual to the frame rule: the invariant  $C_1$  is known inside, unknown outside.

## Type soundness

The type system is proven sound via a standard syntactic argument, which involves subject reduction and progress theorems.

A key lemma is *Revelation*: roughly speaking, a valid type derivation would remain valid if all hidden capabilities were revealed to the world.

### Revelation

A valid judgement remains valid after a previously hidden invariant  ${\cal R}$  is revealed:

Lemma (Revelation)

```
 \begin{array}{ll} \Gamma \vdash v : \tau & \text{implies} & \Gamma \otimes R \vdash v : \tau \otimes R \\ \Gamma; C \vdash t : \sigma & \text{implies} & \Gamma \otimes R; R \wedge (C \otimes R) \vdash t : R \wedge (\sigma \otimes R) \end{array}
```

# Revelation: excerpt of proof

Here is the case of an application:

This is still a valid application, thanks to the equality:

$$(\sigma_1 \to \sigma_2) \otimes R = (R \land (\sigma_1 \otimes R)) \to (R \land (\sigma_2 \otimes R))$$

### How Revelation is used

The gist of the subject reduction proof is that  $anti-frame\ extrudes\ up$  through evaluation contexts:

$$AF = \frac{\frac{\Delta}{\Gamma \otimes R; C \otimes R \vdash t : R \land (\sigma \otimes R)}}{\frac{\Gamma; C \vdash t : \sigma}{\dots}} = \frac{\frac{\Delta}{\Gamma \otimes R; C \otimes R \vdash t : R \land (\sigma \otimes R)}}{\frac{\Gamma' \otimes R; R \land (C' \otimes R) \vdash E[t] : R \land (\sigma' \otimes R)}{\Gamma'; C' \vdash E[t] : \sigma'}} AF$$

The proof is immediate: apply Revelation to (the type derivation for) the evaluation context  $E[\cdot]$ .

### How Revelation is used

This proof technique backs up the intuition that an application of the anti-frame rule amounts to an application of the higher-order frame rule to the evaluation context.

Note: I am quite confident that the type system is sound, but am not done writing the proof yet.

### Contents

- Introduction
- Basics of the type system
- A higher-order anti-frame rule
- Applications
- Conclusion
- Bibliography

# Applications

If there is time, I would like to present three applications of the anti-frame rule:

- untracked references, in the style of ML;
- untracked lazy thunks;
- a generic fixed point combinator.

### Untracked references

In this type system, references are *tracked*: a reference cannot be read or written unless an appropriate capability is presented. This is heavy — capabilities are *linear* — but allows reasoning about *state changes*.

In ML, references are untracked: no capability is required to read or write a cell, and references can be aliased. This is lightweight, but the type of a reference must remain fixed forever.

### Untracked references

Tracked and untracked references have different qualities, so it seems desirable for a programming language to offer both.

But wouldn't that be redundant?

Yes. Type theorists will be happy to hear that, at least in principle, untracked references can be encoded in terms of tracked references and the anti-frame rule.

### Untracked references

The following two slides present the encoding.

For simplicity, the first slide shows integer references. The second slide presents the general case of references to an arbitrary value type a.

## Untracked integer references

```
def type uref =

 a non-linear type!

   (unit \rightarrow int) \times (int \rightarrow unit)
let mkuref : int \rightarrow uref =
\lambda(v: int).
   let \rho, (r : \lceil \rho \rceil) = ref v in
                                                             - got \{ \rho : \text{ ref int } \}
   hide R = \{ \rho : \text{ ref int } \} outside of
   let uget : (R \land unit) \rightarrow (R \land int) =
      \lambda(). get r
   and uset : (R \land int) \rightarrow (R \land unit) =
      \lambda(v : int). set (r, v)
   in (uget, uset)

    this pair has type uref ⊗ R

                                                              - to the outside, uref
```

### Generic untracked references

```
def type uref a =
                                                                       - parameterize over a
   (unit \rightarrow a) \times (a \rightarrow unit)
let mkuref : \forall a.a \rightarrow \text{uref } a =
\lambda(v:a).
   let \rho, (r : \lceil \rho \rceil) = ref v in
                                                                      — got { ρ: ref a }
   hide R = \{ \rho: \text{ ref } a \} \otimes R \text{ outside of } -\text{got } \{ \rho: \text{ ref } a \} \otimes R \}
   let uget : (R \land unit) \rightarrow (R \land (a \otimes R)) = -that is, R
      \lambda(). get r
                                                                       - also \{ \rho : ref (a \otimes R) \}
   and uset: (R \land (a \otimes R)) \rightarrow (R \land unit) =
      \lambda(v:a\otimes R). set (r, v)
   in (uget, uset)
                                                                       - type: (uref a) \otimes R
                                                                       - to the outside, uref a
```

# Lazy thunks

I now define *lazy thunks*, which are built once and can be forced any number of times.

Thunks are untracked and can be freely aliased. Yet, the type system guarantees that each thunk is evaluated at most once.

A thunk contains a hidden reference to an internal state with *three* possible colors (unevaluated, being evaluated, evaluated). Any attempt to ignore the dangers of *re-entrancy* and use only two colors would be ill-typed, by virtue of the anti-frame rule.

## Lazy thunks - part 1

```
def type thunk a =
   unit \rightarrow a
def type state \gamma a =
                                           internal state:
   W(y \wedge unit) + G unit + B a - white/grey/black
let mkthunk : \forall ya.(y \land ((y \land unit) \rightarrow a)) \rightarrow thunk a =
  \lambda(f: (y \land unit) \rightarrow a). - got y
      let \rho, (r : \lceil \rho \rceil) = ref(W()) in -got \{ \rho : ref(state \ \gamma \ a) \}
      hide R = \{ \rho : \text{ ref (state } \gamma \text{ a) } \} \otimes R \text{ outside of }
                                                     - aot R
                                                     - f: ((y \land unit) \rightarrow a) \otimes R
                                                     - f: (R \land (y \otimes R) \land unit) \rightarrow (R \land (a \otimes R))
```

### Lazy thunks - part 2

```
let force : (R \land unit) \rightarrow (R \land a \otimes R) =
  \lambda().
                                            - state \gamma a = W (\gamma \wedge unit) + G unit + B a
     case get r of
                                    - got R = \{ \rho : ref (state \gamma a) \} \otimes R
     | W () \rightarrow
                                    - got \{ \rho: ref (W unit + G ⊥ + B ⊥) \} ∧ (γ ⊗
       set (r, G ());
                                  - got R \wedge (v \otimes R)
        let v : (a \otimes R) = f() in - got R; (v \otimes R) was consumed by f
        set (r. B v);
                                          - aot R
     \mid G \mid G \mid G \rightarrow fail
                                            - without y \otimes R, invoking f is forbidden
     \mid B (v : a \otimes R) \rightarrow v
  in force
                                            - force: (thunk a) \otimes R
                                            - to the outside, thunk a
```

# A fixed point combinator

The fixed point combinator ties a knot in the store in the style of Landin.

It is perhaps not very surprising, but illustrates:

- a use of the anti-frame rule at order 3;
- a delayed initialization, via a strong update;
- a hidden invariant that does not hold upon entry, but does hold upon exit, of the hide construct.

# A fixed point combinator

```
let fix: \forall a_1 a_2 . ((a_1 \rightarrow a_2) \rightarrow (a_1 \rightarrow a_2)) \rightarrow a_1 \rightarrow a_2 =
\lambda(f:(a_1 \rightarrow a_2) \rightarrow (a_1 \rightarrow a_2)).
   let \rho, (r : [\rho]) = ref() in - got { <math>\rho: ref unit }
   hide R = \{ \rho : \text{ ref } (a_1 \rightarrow a_2) \} \otimes R \text{ outside of }
                                                  - haven't got R vet!
   let q:(a_1 \rightarrow a_2) \otimes R = -q invokes !r
      \lambda(x: a_1 \otimes R), aet r x — within a, got R
   in let h: (a_1 \rightarrow a_2) \otimes R = - h invokes f, routing recursive calls to g
    \lambda(x:a_1\otimes R). fax
                                       - f: ((a_1 \rightarrow a_2) \rightarrow (a_1 \rightarrow a_2)) \otimes R
   in set (r, h);
                                                  - a strong update establishes R
   h
                                                  - got R now, as required by anti-frame
                                                  - h: (a_1 \rightarrow a_2) \otimes R
                                                  - to the outside, a_1 \rightarrow a_2
```

### Contents

- Introduction
- Basics of the type system
- A higher-order anti-frame rule
- Applications
- Conclusion
- Bibliography

### Conclusion

In summary, a couple of key ideas are:

- a practical rule for hiding state must be in direct style;
- it is safe for a piece of hidden state to be untracked, as long as its invariant holds at every interaction between Term and Context.

There are more details in the paper [Pottier, 2008].

### Future work

Here are a few directions for future research:

- formally relate frame and anti-frame via a CPS transform;
- extend the functional interpretation developed with Charguéraud in the absence of anti-frame.

# Appendix: typing rules for values

# Appendix: typing rules for terms

$$\begin{array}{c} \text{app} & \text{sub-left} & \text{sub-right} \\ \text{rank} & \Gamma \vdash v : \sigma_1 \to \sigma_2 & \Gamma; C_2 \vdash t : \sigma & \Gamma; C \vdash t : \sigma_1 \\ \hline \Gamma; C \vdash v : C \land \tau & \hline \Gamma; C \vdash (v \, t) : \sigma_2 & \hline \Gamma; C_1 \vdash t : \sigma & \hline \Gamma; C \vdash t : \sigma_2 \\ \hline \exists \rho\text{-elim} & & & \Gamma; C_2 \vdash t : \sigma \\ \hline \Gamma; C \vdash t : \sigma & \rho \# \Gamma, \sigma & \hline \Gamma; C_2 \vdash t : \sigma \\ \hline \Gamma; (\exists \rho.C) \vdash t : \sigma & \hline \Gamma; (C_1 \land C_2) \vdash t : (C_1 \land \sigma) \\ \hline & & & \hline \Gamma; C_2 \vdash t : \sigma \\ \hline \hline \Gamma; C_2 \vdash t : \sigma & \hline \end{array}$$

# Appendix: some subtyping rules

```
func : \tau \equiv \exists \rho . \{\rho : \tau\} [\rho]
```

 $\text{free} \ : \ \textit{C} \leq \textit{\emptyset}$ 

embed-rgn :  $\{\rho_1 : \exists \rho_2.\theta\} \equiv \exists \rho_2.\{\rho_1 : \theta\}$ embed-cap :  $\{\rho_1 : C \land \theta\} \equiv C \land \{\rho_1 : \theta\}$ 

# Appendix: pairs

```
\operatorname{proj}^{1}: \{\rho: t_{1} \times \theta_{2}\} [\rho] \rightarrow \{\rho: t_{1} \times \theta_{2}\} t_{1}
```

 $\text{focus-pair}^1 \ : \ \{\rho:\theta_1\times\theta_2\} \equiv \exists \rho_1.\{\rho:[\rho_1]\times\theta_2\}\{\rho_1:\theta_1\}$ 

# Appendix: sums

```
case : \{\rho: \theta_1 + \theta_2\}([\rho] \times ((\exists \rho_1.\{\rho: [\rho_1] + \bot\}\{\rho_1: \theta_1\}[\rho_1]) \to \sigma) \times ((\exists \rho_2.\{\rho: \bot + [\rho_2]\}\{\rho_2: \theta_2\}[\rho_2]) \to \sigma)) \to \sigma
```

 $\begin{array}{lll} & \text{sub-sum}^1 & : & \left\{\rho:\theta_1+\bot\right\} \leq \left\{\rho:\theta_1+\theta_2\right\} \\ & \text{focus-sum}^1 & : & \left\{\rho:\theta_1+\bot\right\} \equiv \exists \rho_1.\left\{\rho:\left[\rho_1\right]+\bot\right\}\left\{\rho_1:\theta_1\right\} \\ \end{array}$ 

### Contents

- Introduction
- Basics of the type system
- A higher-order anti-frame rule
- Applications
- Conclusion
- Bibliography

(Most titles are clickable links to online versions.)

Birkedal, L., Torp-Smith, N., and Yang, H. 2006.

Semantics of separation-logic typing and higher-order frame rules for Algol-like languages.

Logical Methods in Computer Science 2, 5 (Nov.).

Charguéraud, A. and Pottier, F. 2007.

Functional translation of a calculus of capabilities.

Submitted.

O'Hearn, P., Yang, H., and Reynolds, J. C. 2004.

Separation and information hiding.

In ACM Symposium on Principles of Programming Languages (POPL).

268–280.

### Bibliography]Bibliography



Pottier, F. 2008.

Hiding local state in direct style: a higher-order anti-frame rule. Submitted.