A Practical, Typed Variant Object Model Or, How to Stand On Your Head and Enjoy the View

Pottayil Harisanker Menon, <u>Zachary Palmer</u>, Scott F. Smith, Alexander Rozenshteyn

The Johns Hopkins University

October 22, 2012

#### Record-Based Encodings

• Foundation for traditional OO languages

- Foundation for traditional OO languages
- Easier to type

- Foundation for traditional OO languages
- Easier to type
- Common [Cardelli '84] [Cook '89] ...

- Foundation for traditional OO languages
- Easier to type
- Common [Cardelli '84] [Cook '89] ...
- Variant-Based Encodings

- Foundation for traditional OO languages
- Easier to type
- Common [Cardelli '84] [Cook '89] ...
- Variant-Based Encodings
  - Actor-based languages (Erlang)

- Foundation for traditional OO languages
- Easier to type
- Common [Cardelli '84] [Cook '89] ...
- Variant-Based Encodings
  - Actor-based languages (Erlang)
  - Harder to type

#### Record-Based Object Encoding (Scala) (OCaml) 1 object a { 1 let a = { 2 } 2 }

• Object fields are record fields

• Object fields are record fields

• Methods are fields with functions

- (Scala) (OCaml)
  1 object a {
  2 val v = 5
  3 def mth(x:Int)
  4 :Int = { x+v }
  5 def foo(x:Unit){} 5 foo = fun () -> ()
  6 };
  7 a.mth(3) 7 a.mth a 3
  - Object fields are record fields
  - Methods are fields with functions
  - Invocation projects methods

- (Scala) (OCaml)
  1 object a {
  2 val v = 5
  3 def mth(x:Int)
  4 :Int = { x+v }
  5 def foo(x:Unit){} 5 foo = fun () -> ()
  6 };
  7 a.mth(3) 7 a.mth a 3
  - Object fields are record fields
  - Methods are fields with functions
  - Invocation projects methods
    - We ignore self-hiding for now.

### Duality



# Variant-Based Encoding (Scala) (OCaml) 1 object a { 2 } 2 match msg with 3 ....

(Scala)
1 object a {
2 val v = 5
3 }

|   | (OCa  | am | I) |     |      |       |
|---|-------|----|----|-----|------|-------|
| 1 | let   | v  | =  | ref | 5    | in    |
| 2 | let   | а  | =  | fun | ms   | sg -> |
| 3 | match |    |    | msg | with |       |
| 4 |       |    |    |     |      |       |

• Fields by closure

- - Fields by closure
  - Methods are message handling cases

- (Scala) (OCaml)
  1 object a {
  2 val v = 5
  3 def mth(x:Int)
  4 :Int = { x+v }
  5 def foo(x:Unit){}
  5 x+!self.v
  6 };
  7 a.mth(3) 7 in a ('mth (a,3))
  (OCaml)
  1 let v = ref 5 in
  1 let v = ref 5 in
  2 let a = fun msg ->
  3 match msg with
  1 let v = ref 5 in
  2 let a = fun msg ->
  3 match msg with
  4 i for msg with
  4 i foo () -> ()
  7 a.mth(3)
  - Fields by closure
  - Methods are message handling cases
  - Invocation is just message passing

- (Scala) (OCaml)
  1 object a {
  2 val v = 5
  3 def mth(x:Int)
  4 :Int = { x+v }
  5 def foo(x:Unit){}
  5 x+!self.v
  6 };
  7 a.mth(3) 7 in a ('mth (a,3))
  (OCaml)
  1 let v = ref 5 in
  1 let v = ref 5 in
  2 let a = fun msg ->
  3 match msg with
  1 let v = ref 5 in
  2 let a = fun msg ->
  3 match msg with
  4 | 'mth (self,x) ->
  6 | 'foo () -> ()
  7 a.mth(3)
  - Fields by closure
  - Methods are message handling cases
  - Invocation is just message passing
  - But this doesn't typecheck!

# 1 match v with 2 | 'Odd y -> y mod 2 = 1 3 | 'Dbl x -> x + x

• Typechecking variant destruction is tricky

# 1 match v with 2 | 'Odd y -> y mod 2 = 1 3 | 'Dbl x -> x + x

Typechecking variant destruction is tricky
Most languages (e.g. Caml) fail on unification

# 1 match v with 2 | 'Odd y -> y mod 2 = 1 : int U bool 3 | 'Dbl x -> x + x

- Typechecking variant destruction is tricky
- Most languages (e.g. Caml) fail on unification
- Union types

# 1 match 'Dbl 2 with 2 | 'Odd y -> y mod 2 = 1 : int U bool 3 | 'Dbl x -> x + x

- Typechecking variant destruction is tricky
- Most languages (e.g. Caml) fail on unification
- Union types are insufficient!

# 1 match 'Dbl 2 with 2 | 'Odd y -> y mod 2 = 1 : int! 3 | 'Dbl x -> x + x

- Typechecking variant destruction is tricky
- Most languages (e.g. Caml) fail on unification
- Union types are insufficient!
- Record construction is heterogeneously typed

# 1 match 'Dbl 2 with 2 | 'Odd y -> y mod 2 = 1 : int! 3 | 'Dbl x -> x + x

- Typechecking variant destruction is tricky
- Most languages (e.g. Caml) fail on unification
- Union types are insufficient!
- Record construction is heterogeneously typed
- Variant destruction is not

**Our objective:** a purely type-inferred variant-based object encoding

**Our objective:** a purely type-inferred variant-based object encoding

This can work! We just need...

**Our objective:** a purely type-inferred variant-based object encoding

This can work! We just need...

- A couple new expression forms
- Weakly dependent types
- Precise polymorphism
- A whole-program typechecking pass

**Our objective:** a purely type-inferred variant-based object encoding

This can work! We just need...

- A couple new expression forms
- Weakly dependent types
- Precise polymorphism
- A whole-program typechecking pass

...and then we reap the benefits!

#### How We Get It: TinyBang

# &

#### Onions

(Extensible, type-indexed records)

#### How We Get It: TinyBang

# Onions

&

(Extensible, type-indexed records)



(Functions with built-in patterns)

#### How We Get It: TinyBang



#### TinyBang

1 'dbl x -> x + x

#### • Methods are scapes



1 'dbl x -> x + x

#### • Methods are scapes: functions with patterns

#### TinyBang

 $_{1}$  ('dbl x -> x + x) 'dbl 3

Methods are scapes: functions with patternsInvoke methods by passing messages

#### TinyBang 1 ('dbl x -> x + x) 'dbl 3

Methods are scapes: functions with patterns
Invoke methods by passing first-class messages
# Variant-Based Object Encoding

#### TinyBang

('dbl x -> x + x) ('dbl 3)

- Methods are scapes: functions with patterns
- Invoke methods by passing first-class messages (just labeled data)

- 1 ('dbl x -> x + x) &
- $_{2}$  ('odd y -> y mod 2 == 1)

• Scapes are combined by *onioning* 

- Scapes are combined by onioning
- Application finds match

```
1 (('dbl x -> x + x) &
2 ('odd y -> y mod 2 == 1)) ('dbl 2)
```

```
1 object a {
2  def dbl(x:Int):Int = { x + x }
3  def pos(y:Int):Boolean = { y % 2 == 1 }
4 }
5 a.dbl(2)
```

- Scapes are combined by onioning
- Application finds rightmost match (asymmetric)

- Scapes are combined by *onioning*
- Application finds rightmost match (asymmetric)
- Subsumes case expressions

- (('dbl x -> x + x) &
- 2 ('odd y -> y mod 2 == 1)) ('dbl 2)

- Scapes are combined by *onioning*
- Application finds rightmost match (asymmetric)
- Subsumes case expressions
- Generalizes First-Class Cases [Blume et. al. '06]

1 ('dbl x -> x + x) & 2 ('odd y -> y mod 2 == 1)

('dbl int  $\cup$  'odd int) -> (int  $\cup$  bool)

• Simple union type loses alignment

1 ('dbl x -> x + x) & 2 ('odd y -> y mod 2 == 1)

('dbl int -> int) & ('odd int -> bool)

- Simple union type loses alignment
- Onion type does not

('dbl int -> int) & ('odd int -> bool)

- Simple union type loses alignment
- Onion type does not
- Weakly dependent type

('dbl int -> int) & ('odd int -> bool)

- Simple union type loses alignment
- Onion type does not
- Weakly dependent type
- Relies heavily on polymorphism

#### • Pure variant model: get/set messages

- 1 ('dbl x -> x + x) &
  2 ('odd y -> y mod 2 == 1) &
  3 'Z 5
  - Pure variant model: get/set messages
    Hybrid model: variant methods, record fields

- 1 ('dbl x -> x + x) &
  2 ('odd y -> y mod 2 == 1) &
  3 'Z 5
  - Pure variant model: get/set messages
  - Hybrid model: variant methods, record fields
    Similar to type-indexed rows [Shields, Meijer '01]

- 1 ('dbl x -> x + x) &
  2 ('odd y -> y mod 2 == 1) &
  3 'Z 5
  - Pure variant model: get/set messages
  - Hybrid model: variant methods, record fields
  - Similar to type-indexed rows [Shields, Meijer '01]
  - Labels implicitly create cells

- 4 **in** 0.**Z** 
  - Pure variant model: get/set messages
  - Hybrid model: variant methods, record fields
  - Similar to type-indexed rows [Shields, Meijer '01]
  - Labels implicitly create cells
  - Field access by projection

- $_{4}$  in ('Z z -> z) o
  - Pure variant model: get/set messages
  - Hybrid model: variant methods, record fields
  - Similar to type-indexed rows [Shields, Meijer '01]
  - Labels implicitly create cells
  - Field access by projection/pattern match

- $_{4}$  in ('Z z -> z) o
  - Pure variant model: get/set messages
  - Hybrid model: variant methods, record fields
  - Similar to type-indexed rows [Shields, Meijer '01]
  - Labels implicitly create cells
  - Field access by projection/pattern match
  - But what about self?



• Add 'self to all parameters



Add 'self to all parameters
 & is pattern conjunction



# Add 'self to all parameters & is pattern conjunction

Add 'self to all call sites



- Add 'self to all parameters
   & is pattern conjunction
- Add 'self to all call sites
- Be happy?

#### Naïve Self: Type Problems

```
1 def obj =
2    if something then
3        ('foo _ & 'self s -> s 'bar ()) &
4        ('bar _ -> 1)
5    else
6        ('foo _ & 'self s -> s 'baz ()) &
7        ('baz _ -> 2)
8 in obj 'foo ()
```

#### Naïve Self: Problems

 $\alpha_{\text{SELF}} = (\text{`foo } \& \text{`self } \alpha_1 \rightarrow \text{int}) \& (\text{`bar } - \rightarrow \text{int}) \\ \text{where } \alpha_1 \text{ has `bar} \\ \bigcup \\ (\text{`foo } \& \text{`self } \alpha_2 \rightarrow \text{int}) \& \\ (\text{`baz } - \rightarrow \text{int}) \\ \text{where } \alpha_2 \text{ has `baz} \end{cases}$ 

#### Naïve Self: Problems

$$\begin{array}{l} \alpha_{\mathrm{SELF}} :> & (\texttt{`foo \_ \& `self $\alpha_1$ -> int) \&} \\ & (\texttt{`bar \_ -> int)} \\ & \texttt{where $\alpha_1$ has `bar} \\ \\ \alpha_{\mathrm{SELF}} :> & (\texttt{`foo \_ \& `self $\alpha_2$ -> int) \&} \\ & (\texttt{`baz \_ -> int)} \\ & \texttt{where $\alpha_2$ has `baz} \end{array}$$

• Classic object encodings [Bruce et. al. '98]

Classic object encodings [Bruce et. al. '98]
Type of self is fixed at instantiation

- Classic object encodings [Bruce et. al. '98]
  - Type of self is fixed at instantiation
  - No object extensibility

• Classic object encodings [Bruce et. al. '98]

- Type of self is fixed at instantiation
- No object extensibility

• Extensible Object Calculus [Fisher et. al. '98]

- Classic object encodings [Bruce et. al. '98]
  - Type of self is fixed at instantiation
  - No object extensibility
- Extensible Object Calculus [Fisher et. al. '98]
  - Prototype objects: extensible but not callable

- Classic object encodings [Bruce et. al. '98]
  - Type of self is fixed at instantiation
  - No object extensibility
- Extensible Object Calculus [Fisher et. al. '98]
  - Prototype objects: extensible but not callable
  - Proper objects: callable but not extensible

- Classic object encodings [Bruce et. al. '98]
  - Type of self is fixed at instantiation
  - No object extensibility
- Extensible Object Calculus [Fisher et. al. '98]
  - Prototype objects: extensible but not callable
  - Proper objects: callable but not extensible
  - Prototypes can be *sealed* into proper objects

- Classic object encodings [Bruce et. al. '98]
  - Type of self is fixed at instantiation
  - No object extensibility
- Extensible Object Calculus [Fisher et. al. '98]
  - Prototype objects: extensible but not callable
  - Proper objects: callable but not extensible
  - Prototypes can be *sealed* into proper objects
  - Sealing is permanent

- Classic object encodings [Bruce et. al. '98]
  - Type of self is fixed at instantiation
  - No object extensibility
- Extensible Object Calculus [Fisher et. al. '98]
  - Prototype objects: extensible but not callable
  - Proper objects: callable but not extensible
  - Prototypes can be *sealed* into proper objects
  - Sealing is permanent
  - Sealing is meta-theoretic
# **Self Solutions**

- Classic object encodings [Bruce et. al. '98]
  - Type of self is fixed at instantiation
  - No object extensibility
- Extensible Object Calculus [Fisher et. al. '98]
  - Prototype objects: extensible but not callable
  - Proper objects: callable but not extensible
  - Prototypes can be sealed into proper objects
  - Sealing is permanent
  - Sealing is meta-theoretic

• TinyBang

# **Self Solutions**

- Classic object encodings [Bruce et. al. '98]
  - Type of self is fixed at instantiation
  - No object extensibility
- Extensible Object Calculus [Fisher et. al. '98]
  - Prototype objects: extensible but not callable
  - Proper objects: callable but not extensible
  - Prototypes can be sealed into proper objects
  - Sealing is permanent
  - Sealing is meta-theoretic
- TinyBang
  - Sealing is encodable (no meta-theory)

# **Self Solutions**

- Classic object encodings [Bruce et. al. '98]
  - Type of self is fixed at instantiation
  - No object extensibility
- Extensible Object Calculus [Fisher et. al. '98]
  - Prototype objects: extensible but not callable
  - Proper objects: callable but not extensible
  - Prototypes can be sealed into proper objects
  - Sealing is permanent
  - Sealing is meta-theoretic
- TinyBang
  - Sealing is encodable (no meta-theory)
  - Sealed objects can be **extended** and **resealed**

# Sealing in TinyBang

```
1 def rec seal = obj ->
2   obj &
3   (msg -> obj ('self (seal obj) & msg)) in
4 def point =
5   'x 2 & 'y 4 &
6   ('l1 _ & 'self self -> self.x + self.y) in
7 def sealedPoint = seal point in
8 sealedPoint 'l1 ()
9 ...
```

```
1 . . .
2 def obj = seal (
3 'x 0 &
4 ('inc _ & 'self self ->
5
     self.x = self.x + 1 in self.x) in
6 obj 'inc (); obj 'inc ();
7 def extobj = seal (
8 obj &
9 ('dbl _ & 'self self ->
     self.x = self.x + self.x in self.x)) in
10
11 extobj 'dbl (); extobj 'inc ()
```

x = 0

```
1 . . .
2 def obj = seal (
3 'x 0 &
4 ('inc _ & 'self self ->
5
     self.x = self.x + 1 in self.x) in
6 obj 'inc (); obj 'inc ();
7 def extobj = seal (
8 obj &
9 ('dbl _ & 'self self ->
     self.x = self.x + self.x in self.x)) in
10
11 extobj 'dbl (); extobj 'inc ()
```

x = 1

```
1 . . .
2 def obj = seal (
3 x 0 &
4 ('inc _ & 'self self ->
5
     self.x = self.x + 1 in self.x) in
6 obj 'inc (); obj 'inc ();
7 def extobj = seal (
8 obj &
9 ('dbl _ & 'self self ->
     self.x = self.x + self.x in self.x)) in
10
11 extobj 'dbl (); extobj 'inc ()
```

x = 2

```
1 . . .
2 def obj = seal (
3 'x 0 &
4 ('inc _ & 'self self ->
5
     self.x = self.x + 1 in self.x) in
6 obj 'inc (); obj 'inc ();
7 def extobj = seal (
8 obj &
9 ('dbl _ & 'self self ->
     self.x = self.x + self.x in self.x)) in
10
11 extobj 'dbl (); extobj 'inc ()
```

x = 4

```
1 . . .
2 def obj = seal (
3 'x 0 &
4 ('inc _ & 'self self ->
     self.x = self.x + 1 in self.x)) in
5
6 obj 'inc (); obj 'inc ();
7 def extobj = seal (
8 obj &
9 ('dbl _ & 'self self ->
     self.x = self.x + self.x in self.x)) in
10
11 extobj 'dbl (); extobj 'inc ()
```

x = 5





**'inc ()** 



'self extobj & 'inc ()



'self obj & 'self extobj & 'inc ()



'self obj & 'self extobj & 'inc ()

```
1 def point = seal ('x 0 & 'y 0 &
2 ('l1 _ & 'self self ->
3 self.x + self.y)) in
4 def mixin = ('nearZero _ & 'self self ->
5 (self 'l1 ()) <= 4) in
6 def mixedPoint = seal (point & mixin) in
7 mixedPoint 'nearZero ()
```

#### Mixins

- 1 def point = ... in
- 4 def mixedPoint = seal (point & mixin) in 5 mixedPoint 'nearZero ()

- Mixins
- Higher-order object extension

```
1 def obj = seal (
2 'x 0 & ('inc _ & 'self self ->
3 self.x = self.x + 1 in self.x)) in
4 def obj2 = seal (
5 (obj &. 'x) & 'y 0 &
6 ('inc _ & 'self self ->
7 self.y = self.y + self.x in self.y)) in
8 ...
```

- Mixins
- Higher-order object extension
- Data sharing

```
1 def obj = seal (
2 'x 0 &
3 ('inc n:int & 'self self ->
4 self.x = self.x + n in self.x) &
5 ('inc n:unit & 'self self ->
6 self 'inc 1) in
7 obj ('inc ()); obj ('inc 4)
```

### Mixins

- Higher-order object extension
- Data sharing
- Overloading

etc.

- Mixins
- Higher-order object extension
- Data sharing
- Overloading
- Classes, inheritance, etc.

### • Subtype constraint system

- Subtype constraint system
- Assign each subexpression a type variable

- Subtype constraint system
- Assign each subexpression a type variable
- Derive initial constraint set over expression

- Subtype constraint system
- Assign each subexpression a type variable
- Derive initial constraint set over expression
- Perform knowledge closure on constraints

- Subtype constraint system
- Assign each subexpression a type variable
- Derive initial constraint set over expression
- Perform knowledge closure on constraints
- Check resulting closure for consistency

- Subtype constraint system
- Assign each subexpression a type variable
- Derive initial constraint set over expression
- Perform knowledge closure on constraints
- Check resulting closure for consistency
- Soundness is proven over inference system

### **Constraint Types**

 $int \cup unit$ 

### **Constraint Types**

### $\mathsf{int} \cup \mathsf{unit}$

↕

# $\alpha \setminus \{ \mathsf{int} <: \alpha, \mathsf{unit} <: \alpha \}$

5 + 3















<mark>5 + 3</mark>









<mark>5 + 3</mark>







5 + 3



5 + 3




















$$(x -> x + x)$$
 5



$$(x \rightarrow x + x) 5$$





$$(x \rightarrow x + x) 5$$























 $(x \rightarrow x + x) 5 : int$ 



#### • Let-bound polymorphism

• Type-parametric methods fail

- Type-parametric methods fail
- Local polymorphism

- Let-bound polymorphism
  - Type-parametric methods fail
- Local polymorphism
  - Objects are not local

- Type-parametric methods fail
- Local polymorphism
  - Objects are not local
  - Requires type annotations

- Type-parametric methods fail
- Local polymorphism
  - Objects are not local
  - Requires type annotations
- TinyBang uses call-site polymorphism

- Type-parametric methods fail
- Local polymorphism
  - Objects are not local
  - Requires type annotations
- TinyBang uses call-site polymorphism
  - Each call site is freshly polyinstantiated

- Type-parametric methods fail
- Local polymorphism
  - Objects are not local
  - Requires type annotations
- TinyBang uses call-site polymorphism
  - Each call site is freshly polyinstantiated
  - Recursion reuses variable contours

#### **Polymorphic Application**

**def** id = x -> x **in** (id () & id 1)

#### **Polymorphic Application**

**def** id = x -> x in (id () & id 1)


#### **Polymorphic Application**

**def** id = x -> x **in** (id () & id 1)



#### **Polymorphic Application**

**def** id = x -> x **in** (id () & id 1)



#### **Polymorphic Application**

**def** id = x -> x **in** (id () & id 1)



#### • Aims to infer types for script-like programs

- Aims to infer types for script-like programs
- Uses type information for better performance

- Aims to infer types for script-like programs
- Uses type information for better performance
- Desugars down to TinyBang

- Aims to infer types for script-like programs
- Uses type information for better performance
- Desugars down to TinyBang
- Provides syntax for classes, modules, etc.

- Aims to infer types for script-like programs
- Uses type information for better performance
- Desugars down to TinyBang
- Provides syntax for classes, modules, etc.
- Enough polymorphism for scripting intuitions

- Aims to infer types for script-like programs
- Uses type information for better performance
- Desugars down to TinyBang
- Provides syntax for classes, modules, etc.
- Enough polymorphism for scripting intuitions
- ...without divergence or exponential blow-up

• TinyBang encodes objects as scapes and onions

- TinyBang encodes objects as scapes and onions
- Variant destruction is heterogeneously typed

- TinyBang encodes objects as scapes and onions
- Variant destruction is heterogeneously typed
- (Re)sealing is encodable as a function

- TinyBang encodes objects as scapes and onions
- Variant destruction is heterogeneously typed
- (Re)sealing is encodable as a function
- Flexible OO structures are encodable

- TinyBang encodes objects as scapes and onions
- Variant destruction is heterogeneously typed
- (Re)sealing is encodable as a function
- Flexible OO structures are encodable
- Heavily uses call-site polymorphism model

- TinyBang encodes objects as scapes and onions
- Variant destruction is heterogeneously typed
- (Re)sealing is encodable as a function
- Flexible OO structures are encodable
- Heavily uses call-site polymorphism model
- Requires whole-program typechecking

- TinyBang encodes objects as scapes and onions
- Variant destruction is heterogeneously typed
- (Re)sealing is encodable as a function
- Flexible OO structures are encodable
- Heavily uses call-site polymorphism model
- Requires whole-program typechecking

# Questions?