## Type Abstraction Rules for References: AComparison of Four Which Have Achieved Notoriety

James WilliamO'Tode Jr.

### Abstract

I present four type abstraction rules which have been introduced by various authors to permit polynorphic type safety in the presence of mutable data. Each of the type abstraction rules is discussed in the context of the language in which is was introduced, and the various abstraction rules are compared.

Categories and Subject Descriptions: D3.1 [Programming Languages] – Fornal Definitions and Theory; D3.3 [Programming Languages] – Language Constructs: Implicit Typing; D1.m[Programming Techniques] – Miscellaneous: Polymphic References;

General Terms: Languages, Type Theory, Polymorphism, Reference Values.

Additional Key Words and Phrases: type systems, polynorphic references, nutation, effect systems, type inference, type reconstruction, imperative types, weak polynorphism. Standard ML, FX-89.

## 1 Type Abstraction Rules

The type abstraction rules I consider here are:

- 1. FX-89-pure: expression abstracted must be pure
- 2. Tofte-applicative: one-level store types
- 3. Danas-III: two-level store types
- 4. MacQueen-weak: type variables have strength

## 2 FX-89-pure

Attaches specific side-effect information to all function arrows and enforces the correctness of these effect specifications. The expression which is abstracted with respect to a type variable must have no (immediate) side-effects.

This ought to make it a very restrictive rule, as compared to the others. (Aside from the fact that I expect the checking of the side-effect specifications to disallow more programs.) However, inserting an explicit type abstraction at the appropriate point within the expression might alleviate the problem

### 2.1 FX-89 Language Syntax

 $\pi$  : P ::= Primitive types  $\upsilon$  : U ::= P primitive t

 $\iota$  : I

::=

: U ::= P primitive type I type identifier  $U \rightarrow U$  function

Identi fiers

The type domain U contains the types which are supplied by the programmer in explicit type declarations. The type of a function encodes the type of its argument

<sup>\*</sup>National Science Foundation Graduate Fellow.

This research was supported by the Defense Advanced Research Projects Agency of the Department of Defense and was monitored by the Office of Naval Research under contract number N00014-83-K-0125.

Authors' address: MIT Laboratory for Computer Science, 545 Technology Square, Cambridge, Massachusetts 02139.

E-mail: james@zermatt.lcs.mit.edu

and its result. If the type of the argument is nononorphic, then it may be omitted. The type  $\forall t.v$  represents the type of polynorphic values abstracted over the type parameter t.

In the expression domain, lambda abstracts  ${\tt E}$  over the ordinary variable  ${\tt I}$ .

## 2.2 Deductive System

I present the typing system of IFX as as formal deduction system consisting of a set of type reconstruction rules. The type system contains generic (i.e. general) type variables, and distinguishes between these generic type variables and the type identifiers which appear in user-supplied types. The type systemal so distinguishes between noncomphic and polynorphic types:

The IFX typing rules make use of an important distinction between the Mand Ttype domains. The rules are designed so that Mtypes may be omitted from formal argument type declarations, but Ttypes may not. Thus, the different levels in the type syntax specify the restrictions on the input programs. The use of syntactically-specified restrictions is intended to communicate clearly to the programmer the limitations of the type reconstruction system

### Type Schenes

The IFX type system supports the generic polynor-phism found in ML, as well as the explicit polynorphism found in Reynolds' second-order polynorphic landsda calculus. In order to provide generic polynorphism type schems are defined which represent the generic (i.e. general) type of a variable which is permitted multiple instantiations:

**Definition** (Type Scheme). A type schem  $\eta$  is a term of the form

$$\tilde{\forall} \alpha_1 \dots \alpha_r$$

where  $\alpha_1 \dots \alpha$  are the generic variables of  $\tau \in T$ 

The symbols  $\tilde{\forall}$  and  $\forall$  are distinguished deliberately:  $\tilde{\forall}$  binds the generic type variables of a type scheme, and  $\forall$  binds type variables within a type.

**Definition (Alpha-renaming)**. Types  $\tau$  and  $\tau$  ' are alpha-renamble (written  $\tau \simeq \tau$  ') iff some renaming of type variables bound in  $\tau$  produces  $\tau$  '.

**Definition (Instantiation)**. The type  $\tau$  ' is an instance of the scheme  $\eta = \tilde{\forall} \alpha_1 \dots \alpha_k \tau$  (written  $\eta \succeq \tau$ ) iff there are nonnomorphic  $\mu$  1...  $\mu$  such that  $\tau[\mu_i/\alpha_i] \simeq \tau$ . (The  $\succeq$  relation extends to type schemes by  $\eta \succeq \eta$  ' iff  $\forall \tau : '\eta \succeq \tau \Rightarrow \eta \succeq \tau$ .)

Note that only Mtypes may be substituted to produce instantiations, and that it is assumed that substitution takes place with renaming of any bound type variables to avoid capture. The result of substituting  $\mu$  for t in  $\tau$  will be written  $\tau[\mu/t]$ . The type scheme  $\eta = \tilde{\forall}.\tau$ , having no generic type variables, will occasionally be abbreviated as  $\tau$ .

The inference rules for explicitly typed terms are presented first. At ype assignment A maps each variable in its domain to a type scheme. The notation A x refers to the type assignment A with the assignment for variable x removed.

The notation  $FCN(\tau)$  refers to the free general type variables of  $\tau$ , and  $FTN(\tau)$  to refer to the free type identifiers of  $\tau$ . Similarly, FCN(A) refers to the free general type variables of the type schemes in the assignment A. Also,  $Cen(A, \tau)$  is defined as follows:

**Definition (Generalization).** The generalization of  $\tau$  with respect to A (written  $\text{Gen}(A, \tau)$ ), is the type schem  $\eta = \tilde{\forall} \alpha_i$ .  $\tau$ , where  $\{\alpha_i\} = \text{FGN}(\tau) - \text{FGN}(A)$ .

## Typi ng Rules

The type reconstruction rules of IFX are as follows:

#### ILAMBDA

$$\frac{A_x + (x:\mu) \vdash e:\tau}{A \vdash \text{ (lambda } (x) \ e):\mu \to \tau}$$

$$\begin{array}{ccc}
A \vdash & e : \tau \to \tau_r \\
A \vdash & e_a : \tau_a \\
\hline
A \vdash & (e & e_a) : \tau_r
\end{array}$$

The above rules describe the typing requirements of value abstraction and value application.

The following rules describe the typing requirements of variables and the ML-style generic let construct.

VARI NST

$$\frac{(x:\tilde{\forall}\alpha_i.\,\tau)\in A}{A\vdash x:\tau\,[\mu/\alpha_i]}$$

ILET

$$A \vdash e_b : \tau_b$$

$$A \vdash e_b : \tau_b' \Rightarrow \operatorname{Gen}(A, \tau_b) \succeq \operatorname{Gen}(A, \tau_b')$$

$$A_x + (x : \operatorname{Gen}(A, \tau_b)) \vdash e : \tau$$

$$A \vdash (\operatorname{let}(x e_b) e) : \tau$$

#### Generic let

The HET and VARINST rules provide the ML-style generic let. HET associates a generic type scheme with the let-bound variable, and VARINST permits each occurence of the variable to be independently assigned any instance of its generic type scheme. The convenience of automatic generalization and instantiation are provided by these two rules. In IFX, the typing rules permit this convenience with the caveat that the automatically deduced type parameters be Mtypes.

The typing power of the ILET rule is equivalent to that provided by rewriting the let expression in the usual way, while making use of open and close:

((lambda 
$$(x : \tau) e[(open x)/x])$$
 (close  $\ell$ )).

However, this transformation is not pure syntactic sugar, because it requires  $\tau$ , the explicitly polynorphic type of the bound variable.

## 3 Tofte-applicative

Contaminates all type variables appearing in any type expression at which the ref constructor is instantiated. The contaminated type variables are imperative, the others are applicative. This distinction is maintained by type abstraction, and is enforced at function call boundaries, etc. The abstraction rule does not permit abstractions of expansive expressions with respect to imperative type variables; expansive expressions are let and application expressions.

#### 3. 1 Definitions

The typing system distinguishes between *inperative* and *applicative* type variables:

$$\begin{array}{ll} t & \in \mathit{AppTyVar} \\ u & \in \mathit{ImpTyVar} \\ \alpha & \in \mathit{TyVar} = \mathit{AppTyVar} \cup \mathit{ImpTyVar} \end{array}$$

**Definition (Type Closure).** The type dosume of  $\tau$  with respect to A (written G os  $_A\tau$ ), is the type scheme  $\eta = \tilde{\forall} \alpha_i . \tau$ , where  $\{ \emptyset \}$  =tyvars  $\tau$  - tyvars A

**Definition (Applicative Type Closure).** The applicative type dosume of  $\tau$  with respect to A (written AppClos  $_A\tau$ ), is the type scheme  $\eta = \tilde{\forall}\alpha_i.\tau$ , where  $\{ \varphi \} = \text{apptyvars } \tau - \text{apptyvars } A$ 

When a type scheme is instantiated, only imperative types may be substituted for imperative type variables. An expression is considered to be *expunsive* if its evaluation might expand the domain of the store (i.e., allocate mutable data). The classification adopted in [Tofte87] is that let expressions and applications are expansive, but lambda abstractions and variable accesses are not.

## 3. 2 Typi ng rul es

The reference creation operator **ref** is assigned the imperative type  $\forall u. u \rightarrow u \, ref$ . The rules which provide type abstraction of expansive and non-expansive expressions in the imperative/applicative systemare as follows:

VARINST

LEF Expansi ve

$$A \vdash e_b : \tau_b$$

$$e_b \text{ is expansive.}$$

$$A_x + (x : \text{AppClos}_{A} \tau_b) \vdash e : \tau$$

$$A \vdash (\text{let } (x e_b) e) : \tau$$

LEF Non-expansi ve

$$A \vdash e_b : \tau_b$$

$$e_b \text{ is non-expansive.}$$

$$A_x + (x : G \text{ os } A \tau_b) \vdash e : \tau$$

$$A \vdash (\text{let } (x e_b) e) : \tau$$

## 3.3 Applicative Types and FX 89

In FX 89, type abstraction is permitted only when the side-effect specifications ensure that the polynorphic expression is referentially transparent. [Tofte87] takes a different approach, based on the concept of applicative types. Tofte classifies certain expressions as expansive, and permits type abstraction of these expressions only with respect to applicative type variables. This type abstraction rule permits different type abstractions than does the FX 89 pure-type-abstraction rule, as I will show later. Perhaps the imperative typing discipline can be combined with the type reconstruction system of FX 89.

## 4 Damas-III

Maintains at wolevel version of imperative types, distinguishing those type variables which have been contaminated already from those which will become contaminated by further application. The deductions carry a set of type variables, and so also do any type schemes which are arrows. Types, however, do not carry sets of type variables, nor is there more than a single top-level such set in a type scheme.

#### 4.1 Definitions

The typing system defines schemes to include a set of type variables:

**Definition (Type dosure).** The type closure of  $\eta$  with respect to type assignment A and type variables

 $\Delta$ (written DO os  ${}_{A}\Delta \eta$ ), is the type scheme  $\eta$  ' =  $\forall$   $\alpha_i \cdot \eta$ , where  $\{\alpha_i\}$  =tyvars  $\eta$  - (tyvars  $\Delta$ tyvars  $\Delta$ 

When a type scheme is instantiated, the substitution is used to expand the set of type variables, and the set of type variables may be spuriously expanded as well.

### 4. 2 Typi ng rul es

The reference creation operator **ref** is assigned the imperative type  $\forall t.t \rightarrow t \ ref * \{t\}$ . The rules which provide type abstraction and instantiation are as follows:

DARI IST

$$\frac{(x:\forall \alpha \tau) \in A}{A \vdash x:\tau[\psi \alpha_i] * \phi}$$

DET

$$A\vdash e_b: \eta_b * \Delta$$

$$A_x + (x: \text{IO} \text{ os } A\eta_b) \vdash e: \eta * \Delta$$

$$A\vdash (\text{let } (xe_b) e): \eta * \Delta$$

The rules which describe the typing requirements of value abstraction and value application are as follows:

**ILAMBDA** 

$$\frac{A_x + (x : \tau_a) \vdash \ e : \tau * \ \Delta}{A \vdash \ (\texttt{lambda} \ (x) \ e) : (\overline{\iota} \to \tau * \ \Delta) * \ \phi}$$

DAPPL

$$\begin{array}{cccc} A \vdash & e : (\bar{a} \to \tau_r) * & \Delta \\ \hline A \vdash & e_a : \tau_a * & \Delta \\ \hline A \vdash & (e & e_a) : \tau_r * & \Delta \end{array}$$

## 5 MacQueen-weak

Attaches numbers to type variables which measure their "weakness" (strength). The numbers indicate how many applications must take place before a reference to the type variable might have been created. Abstraction is permitted only with respect to type variables whose weakness remains positive. Wakness is downward contaminating, and the reference constructor is the source of contamination. Afurther restriction is not yet well understood: An instantiation of a let-bound variable is strength-limited somehow, related to the outermost abstraction level at which the expression of which it is part appears as an operand. Better figure this out.

#### 5. 1 Definitions

The typing system distinguishes between *inperative* and applicative type variables:

**Definition (Strength Limit).** The type  $\tau$  is w-strength limited, written  $[\tau]$   $^w$ , iffall type variables in  $\tau$  with non-infinite strength have strength less than or equal to w.

**Definition (Strengthening).** The strengthening of  $\tau$ , written  $[\tau]^{++}$ , is the type in which all type variables with non-infinite strength have incrementally larger strength. So  $[\tau_a \to \tau_b]^{++} \equiv [\tau_a]^{++} \to [\tau_b]^{++}$ , and  $[\alpha^{\infty}]^{++} \equiv \alpha^{\infty}$ , but  $[\alpha^w]^{++} \equiv \alpha^{w+1}$ .

**Definition (Weak Type Closure).** The weak type closure of  $\tau$  with respect to A (written WakO os  $_A\tau$ ), is the type scheme  $\eta = \tilde{\forall}\alpha_i^{w_i}.\tau$ , where  $\{\ \wp\}$  =tyvars  $\tau$  - tyvars A, and  $w_i = min\{\ w|\alpha_i^w \in \text{tyvars }\tau\}$ .

When a type schema is instantiated, the type substituted for a type variable must not be stronger than the type variable.

## Weak Typing Rules

The type reconstruction rules of MacQueen-weak are as follows:

WAMBDA

WAPPL

$$A \vdash e : [\tau \to \tau_r]^{++}$$

$$A \vdash e_a : \tau_a$$

$$\frac{[\tau_a]^0}{A \vdash (e e_a) : \tau_a}$$

The above rules describe the typing requirements of value abstraction and value application.

The following rules describe the typing requirements of variables and the Mrstyle generic let construct.

WNST

$$(x : \tilde{\forall} \alpha_i^{w_i} . \tau) \in A$$

$$\frac{[\tau_i]^{w_i}}{A \vdash x : \tau [y/\alpha_i]}$$

WET

$$\frac{A \vdash e_b : \tau_b}{A \vdash (x : \text{WakCl os} \quad A \tau_b) \vdash e : \tau} \\
\frac{A \vdash (\text{let} (x e_b) e) : \tau}{A \vdash (\text{let} (x e_b) e) : \tau}$$

The reference value constructor **ref** is assigned the type  $\forall \alpha^1. \alpha^1 \rightarrow \alpha^1 ref$ .

## 6 Comparison of Abstraction Rules

## 6. 1 Damas - III > Tofte - applicative

[Tof te87] provides this example on page 73.

Danas-III can type this systembecause the let expression defining f is abstractable with respect to the type of y. This is the case because the two-level analysis of the allocated types of the let expression reveals that none are already allocated, although the type of y will be allocated by further application.

Tofte-applicative cannot type this systembecause the one-level analysis reveals marely that the type of y is-or-will-be allocated, and the let expression is considered expansive, so the type abstraction is not permitted.

## 6.2 Tofte-applicative > Danas-III

(2) No known example.

[Tof te87] states on page 73 that an embedding exists.

## 6.3 MacQueen-weak > Tofte-applicative, Danas-III

[Tofte 87], Example 4.5, mentioned on page 74.

end

MacQueen-weak can type this example, because the counting mathods used by the typing algorithm deduce that fold must be applied three times before any allocation occurs, and since fast\_reverse is defined by applying fold only twice, fast\_reverse still has a type of strength one, and so may be generalized with respect to the type of the elements of the list.

Tofte-applicative cannot type this example because the one-level store typing analysis considers the expression (fold cons []) to be expansive, and therefore does not permit the type abstraction.

Damas-III cannot type this example because the two-level method also considers all type variables to have been allocated by the evaluation of (fold cons []), and therefore does not permit the necessary type abstraction.

# 6.4 Danas-III, Tofte-applicative > Mac Queen-weak

```
(4)    let rid = fn x => !(ref x)
    in let f = fn y => rid (fn a => a) y
        in (f 0;
            f true)
        end
end
```

Tofte-applicative can type this example, because the defining expression for **f** is a lambda abstraction, which is considered non-expansive, and so the type abstraction with respect to the type of **y** is permitted even though it is an imperative type. (Damas-III can also type this

example, because the two-level analysis also reveals that the lambda expression defining f has not yet allocated at any types.)

McQueen-weak cannot type this example, because rid must be given a type with strength one, and yet rid, after instantiation, is applied twice. This lowers the strength so that the strength of the type of y becomes zero. Therefore, the type abstraction is not permitted.

## 6.5 FX-89-pure > Danas-III, Tofte-applicative, MacQueen-weak

FX 89-pure can type this example because the side-effect analysis system correctly determines that nop has no latent side-effects, because the evaluation of nop applied to any arguments f and x will marely return x. If f were applied by nop, then the latent effect of the type of nop would include the latent effect of its argument type, the type of f. Therefore, the defining expression for h is pure, and may be abstracted with respect to the type of b.

Danas-III and Tofte-applicative cannot type this example because the store-typing analysis methods assume that allocation has occurred at the type of a during the evaluation of the defining expression for h (the application of nop). The binding of g in the definition of nop constrains the type of a to be the same as the type of b. Therefore, type abstraction with respect to the type of b is not permitted.

McQueen-weak cannot type this example because the maximumweakness permitted for the type of a is zero, because it is an operand of nop. Therefore, the type of b is forced to strength zero and the type abstraction is not permitted. My intuition for this is that nop is presumed to apply its arguments completely.

# 6.6 Danas-III, Tofte-applicative, MacQueen-weak > FX-89-pure

(6) let  $k = fn a \Rightarrow$ 

```
let r = ref a
        in fn b => !r
        end
in let f = k []
        in (f 0;
        f true;
        false)
    end
end
```

Danas-III, Tofte-applicative, and McQueen-weak can type this example because the defining expression for f can be abstracted with respect to the type of b. All three systems will not permit abstraction with respect to the type of a, because an allocation at that type will have occured. However, this does not prevent the other abstraction, because the types of a and b are not related.

FX 89-pure cannot type this example because the abstraction rule requires that no allocations have taken place, and does not distinguish between the type variables at which allocations have taken place and the type variables at which no allocation have (or will) occur.

## 6.7 Further Speculation

However, the above example will be typed by FX 89-pure if an explicit abstraction is inserted within the definition of k. It is also interesting to observe that even with an explicit type abstraction in the definition of k, it will not be possible to give f a generic type, because of the pure-abstraction rule. Yet f will be automatically projected as required in this example, and f can be opened explicitly as well. Special casing the abstraction rule in let to permit generalization by opening would circumvent this peculiarity, although it will not eliminate the need for explicit abstractions.

Also, I do not expect explicit abstractions to solve this problemin general. Including types in allocation (and perhaps other) effects, and relaxing the pureabstraction rule to examine the side effects and selectively permit type abstractions should provide a much more general treatment of this problem I call this the "Alloc@T" typing system This systemis essentially the systemwhich is mentioned in [Danas85] on pages 90-91, where he observes that attaching sets of types to type arrows will complicate the unification algorithm for types. Danas-III therefore attaches a set of types to type scheme arrows and also a set of types to typing assertions. Tofte-applicative may be viewed as attaching a single set of types to type schemes.

## 7 Summary

Danas-III is strictly superior to Tofte-applicative, but MacQueen-weak and FX-89-pure are incomparable to either of the above. Tofte has suggested in [Tofte89] that MacQueen-weak is strictly superior to Tofte-applicative, but this is not the case (see example (4)).

## Appendix (snh)

Examples provided in snh syntax.

(2) No known example.

```
(5)    let fun id b = b
        and rid a = !(ref a)
        and nop f x =
            let fun g y = f x
            in x
            end
        val h = nop rid id
        in (h 0;
```

```
h true)
         end;
(6)
        let fun ka =
                  let val r = ref a
                   in fn b \Rightarrow !r
                   end
            val f = k []
         in (f 0;
             f true;
             false)
         end;
Appendix (fx)
Examples provided in FX-89 syntax.
(1)
(let ((f (let ((x ((lambda (x) x) 1)))
           (lambda (y) (get (new y))))))
  (begin
    (f 1)
    (f #t)))
(2) No known example.
(3)
(let ((fold (lambda (f i) (lambda (l)
        (let ((data (new 1))
               (result (new i)))
          (begin
             (while (not (null? (get data)))
               (begin
                 (set result (f (car (get data))
                                 (get result)))
                 (set data (cdr (get data)))))
             (get result))))))))
  (let ((fast_reverse (fold cons nil)))
    (begin
      (fast_reverse (list 3 5 7))
      (fast_reverse (list #t #t #f)))))
(4)
        (let ((id (lambda (x) x)))
               (rid (lambda (x) (get (new x)))))
          (let ((f (lambda (y) ((rid id) y))))
             (begin
               (f 0)
               (f #t))))
(5)
(let ((id (lambda (b) b))
```

```
(rid (lambda (a) (get (new a))))
      (nop (lambda (f)
             (lambda (x)
                (let ((g (lambda (y) (f x))))
                 x)))))
  (let ((h ((nop rid) id)))
    (begin
      (h 0)
      (h #t))))
(6)
        (let ((k (lambda (a)
                    (let ((r (new a)))
                      (lambda (b) (get r))))))
          (let ((f (k nil)))
            (begin
              (f 0)
              (f #t)
              #f)))
```

## **References**

- [Dunas 82] Danas, L., Milner, R., "Principal typeschemes for functional programs", Recedings of the 9th Annual Symposium on Rinciples of Rogrammy Languages, January 1982, pages 207-212
- [Danas 85] Danas, L., "Type Assignment in Programming Lanuages", Ph. D. Thesis CST-33-85, University of Edinburgh, April 1985.
- [Gfford87] Gfford, D K, Jouvel ot, P., Lucassen, J. M, Shel don, M A, The FX-87 Reference Manual, MT/LCS/TR-407, October 1987.
- [Hindley69] Hindley, R, "The principal type-scheme of an object in combinatory logic", Transactions of the American Mithematical Society, vol. 146, 1969, pages 29-60.
- [Lucassen87] Lucassen, J. M., Types and Effects: Touards the Integration of Functional and Imperative Programming, Ph. D. Thesis MT/ICS/TR-408, Missachusetts Institute of Technology, September 1987.
- [MacQueen84] MicQueen, D, "Modules for Standard Mi", Proceedings of the 1984 ACMOnference on IISP and Functional Programmy, 1984, pages 198–207.
- [Milner 78] Milner, R, "A Theory of Type Polymrphismin Programming", Journal of Granter and SystemSciences, vol. 17, 1978, pages 349–375.

- [Morris 68] Morris, J. H., *Landa-Gledus Milels of Regionning Languages*, Massachusetts Institute of Technology, MC-TR-57, 1968.
- [O'Toole89] O'Toole, James William, Jr., Type Reconstruction with Erst Class Polymorphic Values, MIT/ICS/TM380, 1989.
- [Tofte87] Tofte, Mids, Operational Semantics and Rolymarphic Type Inference, Ph. D Thesis, University of Elinburgh, 1987.