

EECS 762 - Project 3 Solution

Perry Alexander
The University of Kansas EECS Department
alex@ittc.ku.edu

February 25, 2004

1 Introduction

The objective of Project 3 is to write an interpreter for simply typed lambda calculus with subtyping and records ($\lambda_{<}$) expressions as defined in *Types and Programming Languages* [1], Chapters 15 and 16. In addition, you were to include booleans and the `if` special form. The definition of the abstract syntax provides the following three forms for $\lambda_{<}$ terms, values and types in:

$$\begin{aligned} t &::= x \mid \lambda x : T. t \mid t t \mid \\ &\quad t.l \mid \{l_i = t_i^{i \in 1..n}\} \\ v &::= \lambda x : T. t \mid \text{true} \mid \text{false} \mid \{l_i = v_i^{i \in 1..n}\} \\ T &::= \text{Bool} \mid T \rightarrow T \mid \{l_i : T_i^{i \in 1..n}\} \mid \text{Top} \end{aligned}$$

The definition for call-by-value evaluation provides the following evaluation rules that will define the evaluation function:

$$\frac{t_1 \longrightarrow t'_1}{t_1 t_2 \longrightarrow t'_1 t_2} \text{ E-APP1}$$

$$\frac{t_2 \longrightarrow t'_2}{t_1 t_2 \longrightarrow t_1 t'_2} \text{ E-APP2}$$

$$\frac{}{(\lambda x : T. t_{12}) v_2 \longrightarrow [x \rightarrow v_2] t_{12}} \text{ E-APPABS}$$

$$\frac{\text{if true then } t_2 \text{ else } t_3}{t_2} \text{ E-IFTRUE}$$

$$\frac{\text{if false then } t_2 \text{ else } t_3}{t_3} \text{ E-IFFALSE}$$

$$\frac{t \rightarrow t'_1}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \rightarrow \text{if } t'_1 \text{ then } t_2 \text{ else } t_3} \text{ E-IF}$$

$$\frac{}{\{l_i = v_i^{i \in 1..n}\}.l_j \longrightarrow v_j} \text{ E-PROJRCd}$$

$$\frac{t_1 \longrightarrow t'_1}{t_1.l \longrightarrow t'_1.l} \text{ E-PROJ}$$

Finally, the definition provides the following typing rules that will define the type inference function:

$$\frac{x : T \in \Gamma}{\Gamma \vdash x : T} \text{ T-VAR}$$

$$\frac{\Gamma, x : T_1 \vdash t_2 : T_2}{\Gamma \vdash \lambda x : T_1. t_2 : T_1 \rightarrow T_2} \text{ T-ABS}$$

$$\frac{\Gamma \vdash t_1 : T_{11} \rightarrow T_{12} \quad \Gamma \vdash t_2 : T_{11}}{\Gamma \vdash t_1 t_2 : T_{12}} \text{ T-APP}$$

$$\frac{\Gamma \vdash t_1 : \text{Bool} \quad \Gamma \vdash t_2 : T \quad \Gamma \vdash t_3 : T}{\Gamma \vdash \text{if } t_1 \text{ then } t_2 \text{ else } t_3 : T} \text{ T-IF}$$

$$\frac{}{\text{true} : \text{Bool}}$$

$$\frac{}{\text{false} : \text{Bool}}$$

Our objective is to: (i) define a data structure for representing $\lambda_{<}$ terms embodying the abstract syntax; (ii) a type derivation function for $\lambda_{<}$ terms embodying the type rules; and (iii) an evaluation function for $\lambda_{<}$ terms embodying the evaluation rules.

2 SubtypedLambdaMonad Module

```
module SubtypedLambdaMonad (Term (..))
where
```

The *SubtypedLambdaMonad* module provides the basic definitions for manipulating simply typed lambda calculus expressions with subtyping and records added. As indicated by the module name, a monad is used to implement the evaluation and type checking processes. The data type representing terms and types is first defined, followed by the type checking function, subtyping function, and the evaluation function.

2.1 Data Types

An abstract type T , is defined to represent the two possible types in $\lambda_{<}$. A constant represents the Boolean type while a pair of types represents the \rightarrow type former application.

```
import Data.List
data Retval a =
    Value a | Error String
    deriving (Eq, Show)

instance Monad Retval where
    Error s >>= k = Error s
    Value a >>= k = k a
    return = Value
    fail = Error

data T =
    TyBool |
    TyTop |
    TyArr T T |
    TyTpl [T] |
    TyRec [(String, T)]
    deriving (Eq, Show)

dom :: T → Retval T
dom t =
    case t of
        TyArr d _ → return d
        TyBool → fail "Type Error - Cannot find domain of a boolean type"
        TyRec _ → fail "Type Error - Cannot find domain of a record type"
        TyTpl _ → fail "Type Error - Cannot find the domain of a tuple tye"
        TyTop → fail "Type Error - Cannot find the domain of top"

ran :: T → Retval T
ran t =
    case t of
        TyArr _ r → return r
        TyBool → fail "Type Error - Cannot find range of a boolean type"
        TyRec _ → fail "Type Error - Cannot find range of a record type"
        TyTpl _ → fail "Type Error - Cannot find the range of a tuple tye"
        TyTop → fail "Type Error - Cannot find range of top"

isArr :: T → Bool
isArr (TyArr _ _) = True
isArr _ = False

isRec :: T → Bool
isRec (TyRec _) = True
isRec _ = False

isTpl :: T → Bool
isTpl (TyTpl _) = True
```

isTpl _ = *False*

A constructed type, *Term*, is defined to represent the abstract syntax for $\lambda_{<}$ terms. Note that this definition is identical to that used in $\lambda_{<}$ except *Lambda* includes a type annotation, the *If* construct is defined, and boolean constants have been added. The implementation of the *Term* type is a *Maybe* type that either represents a legal term or a form that cannot be evaluated.

```
data Term =
  TmTrue | TmFalse |
  If Term Term Term |
  Var Int |
  Lambda T Term |
  AppV Term Term |
  AppN Term Term |
  Rec [(String, Term)] |
  Tpl [Term] |
  ProjRcd Term String |
  ProjTpl Term Int
deriving (Eq, Show)
```

The *value* function defines a predicate that specifies $\lambda_{<}$ terms that are values. By definition, *TmTrue*, *TmFalse*, *Lambda* forms, *Rec* and *Tpl* are constructors for values.

```
value :: Term → Bool
value TmTrue = True
value TmFalse = True
value (Lambda _ _) = True
value (Rec _) = True
value (Tpl _) = True
value _ = False
```

The *TmTrue* and *TmFalse* values represent the Boolean values true and false respectively. The *If* constructor defines a classical if-expression. The *Var* constructor identifies the index for a variable and corresponds with the *x* form in the abstract syntax. The *Lambda* constructor defines an abstraction by specifying a term and corresponds with the $\lambda x : T.t$ form in the abstract syntax. The *AppV* constructor defines the application of one term to another using call-by-value semantics. Similarly, the *AppN* constructor defines the application of one term to another using call-by-name semantics. Both forms correspond with the *t t* form in the abstract syntax.

The original λ language implemented untyped lambda calculus. Thus, the maintenance of context information was unnecessary. For $\lambda_{<}$, the types of variables must be maintained as a part of context for type checking. The Γ type is used to store a list of variable bindings that will be used to maintain the types associated with variables in context.

```
type  $\Gamma$  = [T]
```

To manipulate the context, the following functions are defined:

```
addBinding ::  $\Gamma \rightarrow T \rightarrow \Gamma$ 
addBinding  $\gamma$  t = t :  $\gamma$ 
```

```

getTypeFromContext ::  $\Gamma \rightarrow \text{Int} \rightarrow \text{Retval } T$ 
getTypeFromContext c v = if ((length c)  $\geq$  v)
                        then return (c !! v)
                        else fail "Type Error - Variable out of scope"

findField ::  $\Gamma \rightarrow \text{String} \rightarrow T \rightarrow \text{Retval } T$ 
findField  $\gamma$  s1 (TyRec l) = findField2  $\gamma$  s1 l

findField2 ::  $\Gamma \rightarrow \text{String} \rightarrow [(String, T)] \rightarrow \text{Retval } T$ 
findField2 _ _ [] = fail "Type Error - Field name not found"
findField2  $\gamma$  s1 ((s2, t) : l) = if s1  $\equiv$  s2
                                then return t
                                else findField2  $\gamma$  s1 l

findProj ::  $\Gamma \rightarrow \text{Int} \rightarrow T \rightarrow \text{Retval } T$ 
findProj _ i (TyTpl tl) = if i  $\geq$  0  $\wedge$  i < (length tl)
                        then return (tl !! i)
                        else fail "Type Error - Tuple index out of range"

```

2.2 Type Derivation

Type derivation is achieved using the $\text{typeof}_{<}$ function. Each case directly corresponds to one of the typing rules defined for $\lambda_{<}$.

```

typeof_{<} ::  $\Gamma \rightarrow \text{Term} \rightarrow \text{Retval } T$ 
typeof_{<}  $\gamma$  TmTrue = return TyBool
typeof_{<}  $\gamma$  TmFalse = return TyBool
typeof_{<}  $\gamma$  (Rec l) = do tl  $\leftarrow$  mapM ( $\lambda(-, x) \rightarrow (\text{typeof}_{<} \gamma x)$ ) l
                      ; return (TyRec (zip (map ( $\lambda(x, -) \rightarrow x$ ) l) tl))
typeof_{<}  $\gamma$  (Tpl l) = do tl  $\leftarrow$  mapM ( $\lambda x \rightarrow (\text{typeof}_{<} \gamma x)$ ) l
                      ; return (TyTpl tl)
typeof_{<}  $\gamma$  (ProjRcd t s) =
  do rectype  $\leftarrow$  typeof_{<}  $\gamma$  t
  ; if (isRec rectype)
    then findField  $\gamma$  s rectype
    else fail "Type Error - Cannot project a non-record type."
typeof_{<}  $\gamma$  (ProjTpl t i) =
  do tpltype  $\leftarrow$  typeof_{<}  $\gamma$  t
  ; if (isTpl tpltype)
    then findProj  $\gamma$  i tpltype
    else fail "Type Error - Cannot project a non-tuple type."
typeof_{<}  $\gamma$  (If t1 t2 t3) =
  if (typeof_{<}  $\gamma$  t1)  $\equiv$  return TyBool
  then let tt1 = typeof_{<}  $\gamma$  t2; tt2 = typeof_{<}  $\gamma$  t3 in
    if tt1  $\equiv$  tt2 then tt2
    else fail "Type Error - If branches of different types"
  else fail "Type Error - If conditional not boolean"
typeof_{<}  $\gamma$  (Var x) = (getTypeFromContext  $\gamma$  x)
typeof_{<}  $\gamma$  (AppV t1 t2) =
  do tt1  $\leftarrow$  typeof_{<}  $\gamma$  t1

```

```

    ; tt2 ← typeof<. γ t2
    ; dtt1 ← dom tt1
    ; if isArr tt1 then
      if (subtype tt2 dtt1) then return tt2
      else fail "Type Error - Argument type is not a subtype of domain"
      else fail "Type Error - Term is not an abstraction"
  typeof<. γ (AppN t1 t2) =
    do tt1 ← typeof<. γ t1
    ; tt2 ← typeof<. γ t2
    ; dtt1 ← dom tt1
    ; if isArr tt1 then
      if (subtype tt2 dtt1) then return tt2
      else fail "Type Error - Argument type is not a subtype of domain"
      else fail "Type Error - Term is not an abstraction"
  typeof<. γ (Lambda ty t) =
    do tt ← typeof<. (addBinding γ ty) t
    ; return (TyArr ty tt)

```

The *subtype* function defines when one type is a subtype of another. *TyBool* is only a subtype of itself. A function type, *TyArr ty11 ty12* is a subtype of *TyArr ty21 ty22* if *subtype ty21 ty11* and *subtype ty12 ty22*. A record, *TyRec* is a subtype of another if it's list of label, type pairs intersects with the other's label, type pairs. *TyTop* is by definition a supertype of everything while all other subtyping attempts are illegal.

```

subtype :: T → T → Bool
subtype TyBool TyBool = True
subtype (TyArr ty11 ty12) (TyArr ty21 ty22) =
  subtype ty21 ty11 ∧ subtype ty12 ty22
subtype (TyRec l1) (TyRec l2) = (intersect l1 l2) ≡ l1
subtype (TyTpl l1) (TyTpl l2) = (isPrefixOf l1 l2)
subtype _ TyTop = True
subtype _ _ = False

```

2.3 Shifting and Substituting

Our implementation uses de Bruijn indices as a basis for representation and evaluation. Thus, definitions of *shift* and *subst* are required to implement elements of the evaluation function. These functions are largely the same as those used in Project 1 except they must deal with type information.

The *shift* definition provides a case for shifting over each valid λ form as defined in the *Term* data type. This definition follows directly from the standard definition of shift from TPL Chapter 8:

```

shift :: Term → Int → Int → Term
shift TmTrue c d = TmTrue
shift TmFalse c d = TmFalse
shift (Rec l) c d =
  (Rec (map (λ(str, t) → (str, (shift t c d))) l))
shift (Tpl l) c d =
  (Tpl (map (λt → (shift t c d)) l))
shift (ProjRcd r s) c d = (ProjRcd (shift r c d) s)
shift (ProjTpl t i) c d = (ProjTpl (shift t c d) i)
shift (Var x) c d = if x < c then (Var x) else (Var (x + d))

```

$$\begin{aligned}
\text{shift } (\text{Lambda } ty \ t) \ c \ d &= (\text{Lambda } ty \ (\text{shift } t \ (c + 1) \ d)) \\
\text{shift } (\text{AppV } t_1 \ t_2) \ c \ d &= (\text{AppV } (\text{shift } t_1 \ c \ d) \ (\text{shift } t_2 \ c \ d)) \\
\text{shift } (\text{AppN } t_1 \ t_2) \ c \ d &= (\text{AppN } (\text{shift } t_1 \ c \ d) \ (\text{shift } t_2 \ c \ d))
\end{aligned}$$

Lambda, *AppV*, and *AppN* terms are shifted by shifting their identified terms. Shifting a *Var* term requires application of the definition from Chapter 8. Specifically, the index for a variable is shifted by d if its index is greater than c .

The *subst* definition again provides a case for substitution over each value λ as defined in the *Term* data type. This definition follows directly from the standard definition of substitution from TPL Chapter 8:

```

subst :: Int → Term → Term → Term
subst j s (Var x) = if x ≡ j then s else (Var x)
subst j s (Lambda ty t) = (Lambda ty (subst (j + 1) (shift s 0 1) t))
subst j s (AppV t1 t2) = (AppV (subst j s t1) (subst j s t2))
subst j s (AppN t1 t2) = (AppN (subst j s t1) (subst j s t2))
subst _ _ TmTrue = TmTrue
subst _ _ TmFalse = TmFalse
subst j s (Rec l) =
  (Rec (map (λ(str, t) → (str, (subst j s t))) l))
subst j s (Tpl l) =
  (Tpl (map (λt → (subst j s t)) l))
subst j s (ProjRcd r str) =
  (ProjRcd (subst j s r) str)
subst j s (ProjTpl t i) =
  (ProjTpl (subst j s t) i)

```

2.4 Call By Value Evaluation Function

The $\text{eval}_{<}$ function provides a standard definition of call-by-value evaluation following from TPL Chapter 5. The function definition is split up into cases corresponding to the evaluation rules. Note that only *AppN* and *AppV* forms can be evaluated. *Lambda* forms are values and *Var* forms are not closed. Although passing a context is allowed in this function, in this project we are concerned only with close terms. Thus, the context can be safely ignored in all evaluation cases.

The form of the evaluation function is a mapping from Γ and *Term* to another *Term*:

```

expr :: Term → Term
expr (Lambda ty t) = t

proj :: [(String, Term)] → String → Retval Term
proj [] _ = fail "Evaluation Error - Field not defined"
proj ((s1, v) : l) s2 = if s1 ≡ s2
  then return v
  else (proj l s2)

eval< :: Term → Retval Term

eval< TmTrue = return TmTrue
eval< TmFalse = return TmFalse
eval< (Rec x) = return (Rec x)

```

```

eval<: (Tpl x) = return (Tpl x)

eval<: (AppV t1 t2) =
  do nt1 ← eval<: t1
    ; nt2 ← eval<: t2
    ; eval<: (shift 0 (shift nt2 0 1) (expr nt1)) 0 (-1))

eval<: (AppN t1 t2) =
  do nt1 ← eval<: t1
    ; eval<: (shift 0 (shift t2 0 1) (expr nt1)) 0 (-1))

eval<: (Lambda ty1 t1) = return (Lambda ty1 t1)

eval<: (Var t1) = fail "Evaluation Error - Cannot evaluate free variable"

eval<: (If TmTrue t2 t3) =
  do tr ← eval<: t2
    ; return tr
eval<: (If TmFalse t2 t3) =
  do fl ← eval<: t3
    ; return fl
eval<: (If t1 t2 t3) =
  do cond ← (eval<: t1)
    ; eval<: (If cond t2 t3)
eval<: (ProjRcd (Rec r) l) = (proj r l)
eval<: (ProjTpl (Tpl t) i) = return (t !! i)

```

3 Type Checking and Evaluation

The objective of type checking is to statically predict the runtime behavior of a code element with respect to the kinds of values produced and expected. Specifically, does a program element always product the kind of value expected? The `typeof` function provides a capability for generating the type associated with a $\lambda_{<}$ term. If such a type exists, then the term is well-typed and should be executed.

To combine type checking and evaluation is a simple matter of: (i) determining the type of a term; and (ii) executing the term if the type exists. This process can be specified using the following template:

```

interpTemplate :: (Term → Retval T) →
  (Term → Retval Term) →
  Term →
  Retval Term
interpTemplate typeof<: eval<: term =
  do termtype ← (typeof<: term)
    ; term' ← (eval<: term)
    ; return term'

evalTemplateStar :: (Term → Retval Term) → Term → Retval Term
evalTemplateStar eval<: term =
  do term' ← (eval<: term)
    ; if (value term')

```



```

then return term'
else (evalTemplateStar eval<:) term'

```

```

interpret :: Term → Retval Term
interpret = interpTemplate (typeof<: []) eval<:

```

4 Testing and Evaluation

To test the interpreter, some functions from the book are provided here. They include the identity combinator, Church Boolean functions, some Church Numbers and the successor function defined for Church numbers.

Note that with the introduction of types, all parameters must have associated type annotations. Although significant static checking results, the flexibility of these functions is diminished. Specifically, for the **ident** combinator a new combinator must be written for each type. No polymorphism exists and the type system is strict, so there is no way to reuse the **ident** combinator definition. Such strictness is common in older programming languages, but new polymorphism implementations render the approach obsolete.

4.1 Identity Combinator

Redefine the *ident* combinator to operate over *Top* to take advantage of subtyping. Evaluating the (*App ident t*) combinator on term defined in $\lambda_{<:}$ should result in *t* regardless of its type:

```

ident :: Term
ident = (Lambda TyTop (Var 0))

testIdent :: [Retval Term]
testIdent =
  map interpret [ident,
    (Rec [("1", TmTrue), ("2", TmFalse), ("3", ident)]),
    (ProjRcd (Rec [("1", TmTrue), ("2", TmFalse), ("3", ident)]) "1")]

testRecord :: [Retval Term]
testRecord =
  let r = (Rec [("1", TmTrue), ("2", TmFalse), ("3", ident)]) in
    map interpret
      [(ProjRcd r "1"),
       (ProjRcd r "3"),
       (ProjRcd r "4"),
       (AppV (ProjRcd r "3") r)]

testRecSubtype :: [Retval Term]
testRecSubtype =
  let l = (Lambda (TyRec [("1", TyBool), ("2", TyBool), ("3", TyBool)])
    (ProjRcd (Var 0) "2")) in
    map interpret
      [(AppV l (Rec [("1", TmTrue), ("2", TmFalse), ("3", TmTrue)])),
       (AppV l (Rec [("1", TmTrue), ("2", TmFalse)])),
       (AppV l (Rec [("2", TmFalse), ("3", TmTrue)]))]

```

$(AppV\ l\ (Rec\ [(\text{"1"}, TmTrue), (\text{"4"}, TmFalse), (\text{"3"}, TmTrue)]))]$

```
testTuple :: [Retval Term]
testTuple =
  let r = (Tpl [TmTrue, TmFalse, ident]) in
    map interpret
      [(ProjTpl r 0),
       (ProjTpl r 2),
       (ProjTpl r 3),
       (AppV (ProjTpl r 2) r)]

testTupleSubtype :: [Retval Term]
testTupleSubtype =
  let l = (Lambda (TyTpl [TyBool, TyBool, TyBool])
            (ProjTpl (Var 0) 1)) in
    map interpret
      [(AppV l (Tpl [TmTrue, TmFalse, TmTrue])),
       (AppV l (Tpl [TmTrue, TmFalse])),
       (AppV l (Tpl [TmFalse, TmTrue])),
       (AppV l (Tpl [TmTrue, ident, TmTrue]))]
```

4.2 Church Boolean Definitions

The boolean combinators must have typed parameters. Real booleans are chosen for simplicity. Church Booleans don't make a great deal of sense in a strongly typed language like this.

```
tru :: Term
tru = (Lambda TyTop (Lambda TyTop (Var 1)))
fls :: Term
fls = (Lambda TyTop (Lambda TyTop (Var 0)))

testTru :: Retval Term
testTru = interpret (AppV (AppV tru ident) fls)

testFls :: Retval Term
testFls = interpret (AppV (AppV fls fls) ident)
```

4.3 Church Number Definitions

The number combinators must have typed parameters. Real booleans are chosen for simplicity. In a word, these definitions are not appropriate anymore, but they are retained for testing.

```
c0 :: Term
c0 = (Lambda TyBool (Lambda TyBool (Var 0)))

c1 :: Term
c1 = (Lambda TyBool (Lambda TyBool (AppV (Var 1) (Var 0))))

c2 :: Term
```

$$c_2 = (\text{Lambda } \text{TyBool } (\text{Lambda } \text{TyBool } (\text{AppV } (\text{Var } 1) (\text{AppV } (\text{Var } 1) (\text{Var } 0)))))$$

$$\begin{aligned} scc &:: \text{Term} \\ scc &= (\text{Lambda } \text{TyBool} \\ &\quad (\text{Lambda } \text{TyBool} \\ &\quad\quad (\text{Lambda } \text{TyBool } (\text{AppV } (\text{Var } 1) \\ &\quad\quad\quad (\text{AppV } (\text{Var } 2) \\ &\quad\quad\quad\quad (\text{AppV } (\text{Var } 1) (\text{Var } 0))))) \end{aligned}$$

4.4 Omega

With the introduction of types, the omega combinator must type its parameter. Boolean is selected here for the sake of simplicity, but other types could be used.

$$\begin{aligned} \Omega &:: \text{Term} \\ \Omega &= (\text{Lambda } \text{TyTop } (\text{AppV } (\text{Var } 0) (\text{Var } 0))) \end{aligned}$$

5 Notes

The evaluation function is built to handle both call-by-value and call-by-name function application. This may end up causing problems if the arguments to record and tuple references are treated as types in the language rather than Haskell string and integer types.

Are record and tuple projections too aggressive for call-by-name application? They are not true functions in the language due to the types of their reference values.

References

- [1] B. Pierce. *Types and Programming Languages*. MIT Press, Cambridge, MA, 2002.