1 Hacking with the untyped call-by-value lambda calculus

Pierce explains Church encodings for booleans, numerals, and operations on them in (TAPL book s. 5.2 p. 58). We would like to define some more advanced functions using only the basic operations scc, plus, prd, times, iszro, test, fix and the constants. The complete list of predefined operations can be found in the appendix, and only these operations can be used in the exercise. Define the following operations on non-negative integers:

1. The greater equal operation $geq (\geq)$

```
geq = \lambda x. \lambda y. iszro (x prd y)
```

2. The greater than operation gr(>)

gr =
$$\lambda$$
x. λ y. (iszro (y prd x)) fls tru

3. The modulo operation mod (e.g. mod c_5 $c_3 = c_2$)

```
mod =
  fix \lambdame. \lambdax. \lambday.
  test (geq x y)
  (\lambdathen. (me (y prd x) y))
  (\lambdaelse. x)
```

4. The Ackermann function *ack* using the basic operations and operations defined in this exercise. The Ackermann function is defined as follows:

$$A(m,n) = \begin{cases} n+1 & \text{if } m = 0\\ A(m-1,1) & \text{if } m > 0 \text{ and } n = 0\\ A(m-1,A(m,n-1)) & \text{if } m > 0 \text{ and } n > 0. \end{cases}$$

2 Uniqueness of terms after reductions

For this exercise assume we will be working with a simple language defined using the following algebraic datatype in Scala:

abstract class Base
case object Empty extends Base
case class Unary(i: Base) extends Base
case class Binary(i: Base, j: Base) extends Base

Let \rightarrow^{β} be a reduction defined by the following reductions:

Binary(Empty, x)
$$\rightarrow$$
 x (β_1)

$$Binary(x, y) \rightarrow Binary(y, x)$$
 (β_2)

and congruence rules:

$$\frac{\mathtt{x} \to \mathtt{x'}}{\mathtt{Unary}(\mathtt{x}) \to \mathtt{Unary}(\mathtt{x'})} \tag{β_3}$$

$$\frac{\mathtt{x} \to \mathtt{x'}}{\mathtt{Binary}(\mathtt{x}, \ \mathtt{y}) \to \mathtt{Binary}(\mathtt{x'}, \ \mathtt{y})} \tag{\beta_4}$$

$$\frac{\mathtt{y}\to\mathtt{y'}}{\mathtt{Binary}(\mathtt{x},\ \mathtt{y})\to\mathtt{Binary}(\mathtt{x},\ \mathtt{y'})} \tag{β_5}$$

Prove that for any elements u, v, w of type Base we have

$$u \to^{\beta} v \land u \to^{\beta} w \Rightarrow \exists x.(v \to^{*\beta} x \land w \to^{*\beta} x)$$

where $\to^{*\beta}$ refers to 0 or more \to^{β} reductions. Your solution has to have a clear explanation for each step in your proof.

Hint: We suggest proving the above statement using induction on the structure of u.

Solution: If v = w then we are done since x = v = w. Therefore it remains to prove the case when $v \neq w$. We prove it using induction on the structure of u.

- Case u = Empty. No reduction applies. Done.
- Case $u = Unary(u_1)$. Then $u \to^{\beta_3} v \land u \to^{\beta_3} w$. Additionally v and w are of the form $Unary(v_1)$ and $Unary(w_1)$, respectively, because of $u_1 \to^{\beta} v_1$ and $u_1 \to^{\beta} w_1$. By IH we have that $\exists x_1.(v_1 \to^{\beta*} x_1 \land w_1 \to^{\beta*} x_1)$. Hence, for $x = Unary(x_1)$, $v \to^{\beta*} x \land w \to^{\beta*} x$.
- Case $u = Binary(u_1, u_2)$.
 - Case $\underline{v = Binary(u_2, u_1)}$. $v \to^{\beta_2} u$ and $u \to^{\beta_w} w$ where β_w corresponds to the specific reduction that lead to w. Hence there exists an x = w because $v \to^{\beta_2} u \to^{\beta_w} x$.
 - Case $w = Binary(u_2, u_1)$. Analogous to the case above, except that x = v.

- Case $u_1 \neq Empty$ where $u_1 \rightarrow^{\beta} u_1'$ and $v = Binary(u_1', u_2)$. Then:
 - * If $w = Binary(u_1'', u_2)$ where $u_1 \to^{\beta} u_1''$. By IH $\exists u_1'''.(u_1' \to^{\beta *} u_1''' \land u_1'' \to^{\beta *} u_1''')$. Since $x = Binary(u_1''', u_2)$, done.
 - * If $w = Binary(u_1, u_2')$ where $u_2 \to^{\beta} u_2'$. Then $v \to^{\beta_w} Binary(u_1', u_2') \land w \to^{\beta_v} x$ where $x = Binary(u_1', u_2')$.
- Case $u_1 = Empty$. From the assumption we know that $v \neq w$. Therefore we pick v = u_2 by $\overline{(\beta_1)}$. Then the only case not covered for w is for step $w = Binary(Empty, u_2')$ where $u_2 \to^{\beta_w} u_2'$. Then $v = u_2 \to^{\beta_w} u_2' = x$ and $w = Binary(Empty, u_2') \to^{\beta_1} u_2' = x$. Done.

q.e.d

3 The call-by-value simply typed lambda calculus with returns

Consider a variant of the call-by-value simply typed lambda calculus specified in the appendix extended to support a new language construct: $return\ t$, which immediately returns a given term t from an enclosing lambda, disregarding any potential further computation typically needed for call-by-value evaluation rules.

The grammar of the extension is defined as follows. We distinguish top-level terms (tt) and nested terms (nt) to make sure that **return** t can only appear inside lambdas:

```
\begin{array}{lllll} v & ::= & \lambda x : \ {\tt T} \cdot nt \mid bv & (values) \\ bv & ::= & {\tt true} \mid {\tt false} & (boolean\ values) \\ nt & ::= & x \mid v \mid nt\ nt \mid {\tt return}\ nt & (nested\ terms) \\ tt & ::= & x \mid v \mid tt\ tt & (top-level\ terms) \\ t & ::= & nt \mid tt & (terms) \\ p & ::= & tt & (programs) \\ T & ::= & {\tt Bool} \mid {\tt T} \rightarrow {\tt T} & (types) \end{array}
```

In this exercise, you are to adjust the existing evaluation and typing rules, so that they correctly and comprehensively describe the semantics of the extension. More precisely, your task is two-fold:

- 1. Extend the evaluation rules (by adding new rules and/or changing existing ones) to express the early return semantics provided by return. Identify the evaluation strategy used by the specification and make sure that your extension adheres to it.
- 2. Extend the typing rules (by adding new rules and/or changing existing ones) to guarantee that types of values returned via return and via normal means are coherent. Make sure that progress and preservation conditions hold for your extension (you don't have to prove that formally, but your grade will be reduced if your extension ends up being unsafe).

Hint: In addition to the immediate type of a term, you also need to keep track of the types of returned terms inside that term. For example, instead of the regular typing judgment $\Gamma \vdash t : T$, you can use the $\Gamma \vdash t : T \mid R$, where R is a set of types of terms, i.e. $\{T_1, ... T_n\}$, that can be returned from t.

Before you begin, think carefully about the following simple term: λx : Bool. (return true) x. Intuitively, it makes sense. Once this lambda is applied, it is going to evaluate to true, regardless of the input. Now, which typing rules would be used to type this term, so that it is accepted by our language? In particular, what type or types need to be assigned to return true?

Solution:

Evaluation rules:

$$\frac{t_1 \longrightarrow t_1'}{t_1 \ t_2 \longrightarrow t_1' \ t_2} \tag{E-APP1}$$

$$\frac{t_2 \longrightarrow t_2'}{v_1 \ t_2 \longrightarrow v_1 \ t_2'} \tag{E-App2}$$

$$\frac{[x \to v_2]t_1 \longrightarrow^* \text{return } v_3}{(\lambda x \colon \mathsf{T}_1. \ t_1) \ v_2 \longrightarrow v_3} \tag{E-AppAbs1}$$

$$\frac{[x \to v_2]t_1 \longrightarrow^* v_3}{(\lambda x \colon \mathsf{T}_1 \cdot t_1) \ v_2 \longrightarrow v_3} \tag{E-AppAbs2}$$

$$\frac{t \longrightarrow t'}{\text{return } t \longrightarrow \text{return } t'} \tag{E-Ret}$$

$$(\text{return } v_1) \ t_2 \longrightarrow v_1 \tag{E-AppRet1}$$

$$t_1 \text{ (return } v_2) \longrightarrow v_2$$
 (E-AppRet2)

$$\texttt{return} \; (\texttt{return} \; v) \longrightarrow v \tag{E-RetRet}$$

Typing rules:

$$\frac{x : T \in \Gamma}{\Gamma \vdash x : T \mid \emptyset}$$
 (T-VAR)

$$\frac{\Gamma,\ t:\mathtt{T}\ |\mathtt{R}}{\forall\ \mathtt{A}.\ \Gamma\vdash\mathtt{return}\ t:\mathtt{A}\mid \{\mathtt{T}\}\cup\mathtt{R}} \tag{T-Ret}$$

$$\frac{\Gamma, \ x : \mathsf{T}_1 \vdash t_2 : \mathsf{T}_2 \mid \emptyset}{\Gamma \vdash (\lambda x \colon \mathsf{T}_1 \cdot t_2) : \mathsf{T}_1 \to \mathsf{T}_2 \mid \emptyset}$$
 (T-Abs1)

$$\frac{\Gamma, \ x : \mathsf{T}_1 \vdash t_2 : \mathsf{T}_2 \mid \{\mathsf{T}_2\}}{\Gamma \vdash (\lambda x \colon \mathsf{T}_1 \cdot t_2) : \mathsf{T}_1 \to \mathsf{T}_2 \mid \emptyset}$$
 (T-Abs2)

$$\frac{\Gamma \vdash t_1 : \mathsf{T}_1 \to \mathsf{T}_2 \mid \mathsf{R}_1 \quad \Gamma \vdash t_2 : \mathsf{T}_1 \mid \mathsf{R}_2}{\Gamma \vdash t_1 \ t_2 : \mathsf{T}_2 \mid \mathsf{R}_1 \cup \mathsf{R}_2} \tag{T-App)}$$

$$\frac{}{\Gamma \vdash \mathsf{true} : \mathsf{Bool} \mid \emptyset} \tag{T-FALSE}$$

$$\frac{}{\Gamma \vdash \mathsf{false} : \mathsf{Bool} \mid \emptyset} \tag{T-True}$$

4 Closed terms

We recall that a term t is closed if it contains no free variables. With that definition in mind prove the following property for the call-by-value untyped lambda calculus (for reference provided in Appendix 1).

Theorem: If t is closed, and $t \longrightarrow t'$ then t' is closed as well.

Note: Remember to state clearly all the steps of your proof, including proofs of any lemmas that you use.

Solution:

We prove the Theorem by induction on the structure of t.

- 1. Let t be a variable x. The solution is immediate since t is closed and the case cannot occur.
- 2. Let t be an abstraction $\lambda x.t_1$. Since $\lambda x.t_1 \not\longrightarrow$, the solution is immediate.
- 3. Let t be an application t_1 t_2 . Then we can have three different cases, based on the used reduction rule for $t \longrightarrow t'$:
 - (a) E-APP1 then $t_1 \longrightarrow t'_1$. As t is closed, then both t_1 and t_2 are closed as well. By induction t'_1 is closed as well. Therefore t'_1t_2 is closed as well.
 - (b) E-APP2 then $t_2 \longrightarrow t'_2$. As t is closed, then both v_1 and t_2 are closed as well. By induction t'_2 is closed as well. Therefore $v_1t'_2$ is closed as well.
 - (c) E-APPABS then $[x \to v_2]t_{12}$ for $(\lambda x.t_{12})v_2$. As t is closed, then both $\lambda x.t_{12}$ and v_2 are closed as well. In order to prove this, we need another lemma which says that substitution preserves the fact that terms are closed (stated and proven below). Done.

Lemma: For $\lambda x.t_1$ and t_2 that are closed, $[x \to t_2]t_1$ is closed as well. Proof by induction on the structure of the lambda term t_1 .

- 1. Let t_1 be a variable y. If x = y, then substitution results in t_2 , which is closed. If $x \neq y$ then substitution results in t_1 . Since from the assumption we know that t_1 is closed, therefore y was not free.
- 2. Let t_1 be an abstraction $\lambda y.t'_1$. If x = y, then the result is immediate, since x is bound in the abstraction, the t_1 is closed. If $x \neq y$, then by induction $[x \to t_2]t'_1$ is closed as well. Therefore $\lambda y.t'_1$ is closed.
- 3. Let t_1 be an application $t_1't_2'$. Since t_1' and t_2' are closed, by induction $[x \to t_2]t_1'$ and $[x \to t_2]t_2'$ are closed as well. Therefore $[x \to t_2](t_1' t_2')$ is closed.

5 Dynamic

In this exercise we extended the Simply Typed Lambda Calculus with a Dynamic type. The Dynamic type allows us to escape the static nature of STLC and create some interesting constructs.

For instance, for STLC with arithmetic expressions, $(\lambda x : Nat.0)(\lambda y : Nat.0)$ is not typeable even though variable x is not used within the body of the value abstraction. With Dynamic, we would be able to write a term like $(\lambda x : Dynamic. 0)$ $(dynamic (\lambda y : Nat. 0) \text{ as } Nat \rightarrow Nat)$ though.

The dynamic construct would therefore *package* a value and its type. *Unpacking* dynamic values is only possible through the typecase construct that matches on the underlying type of Dynamic (similarly to Sums which were defined during the lectures).

Below we give a tentative formalization of this extension.

Typing and evaluation rules:

For this exercise you have to:

- Show that progress and preservation **do not** hold by giving a brief example of each violation.
- Make the extension sound (by making minimal number of changes to the grammar, typing and evaluation rules). Note that having a reduction rule which reduces a term to itself is not a desirable semantics for this exercise.

• Prove progress of your extension by detailing new cases for the inductive proof. You need to discuss only the cases that result from the evaluation and typing rules that were introduced, and, if needed, you may use (without proving them) the standard lemmas of inversion of typing, canonical forms, and uniqueness of types.

Although you do not need to prove preservation you have to make sure that your changes in the previous step make it succeed.

Solution:

Progress doesn't hold. One example is a missing case for a Dynamic type tag in a typecase construct: typecase (dynamic $\lambda x: Nat.0$ as $Nat \to Nat$) match {case $x: Nat \Rightarrow x$ } Preservation doesn't hold because we miss a necessary premise in (T-Dynamic) typing rule: typecase (dynamic $\lambda x: Nat.0$ as Nat) match case $x: Nat \Rightarrow x$ and before the reduction step the type of the expression is Nat and afterwords it is Nat \to Nat.

Correct typing and evaluation rules:

$$\frac{\Gamma \vdash \mathsf{t} : \mathsf{T}}{\Gamma \vdash \mathsf{dynamic} \ \mathsf{t} \ \mathsf{as} \ \mathsf{T} : \mathsf{Dynamic}} \qquad (\mathsf{T}\text{-}\mathsf{Dynamic})$$

$$\frac{\Gamma \vdash \mathsf{t} : \mathsf{Dynamic}}{\Gamma, x_i : T_i \vdash \mathsf{t}_i : \mathsf{T}}$$

$$\frac{\Gamma \vdash \mathsf{t} : \mathsf{dynamic}}{\Gamma \vdash \mathsf{t} : \mathsf{dlse} \ \mathsf{T}} \qquad (\mathsf{T}\text{-}\mathsf{TYPECASE})$$

$$\frac{k \in 1... \land T_k = \mathsf{T}}{\Gamma \vdash \mathsf{telse} \ \mathsf{t} \ \mathsf{match} \ \{\mathsf{case} \ \mathsf{x}_i : \mathsf{T}_i \ \Rightarrow \ \mathsf{t}_i \ \mathsf{else} \ \mathsf{telse}\} : \mathsf{T}} \qquad (\mathsf{T}\text{-}\mathsf{TYPECASE})$$

$$\frac{k \in 1... \land T_k = \mathsf{T}}{\mathsf{for} \ \mathsf{each} \ \mathsf{j}, \mathsf{j} < k \ \Rightarrow \ \mathsf{T}_j \neq \mathsf{T}} \qquad (\mathsf{E}\text{-}\mathsf{TYPECASE})$$

$$\mathsf{typecase} \ (\mathsf{dynamic} \ \mathsf{v} \ \mathsf{as} \ \mathsf{T}) \ \mathsf{match} \ \{\mathsf{case} \ \mathsf{x}_i^{\mathsf{i}\in 1..n} : \mathsf{T}_i \ \Rightarrow \ \mathsf{t}_i \ \mathsf{else} \ \mathsf{t}_{else}\} \rightarrow \mathsf{t}_{else} \qquad (\mathsf{E}\text{-}\mathsf{TYPECASE}\text{-}\mathsf{DYNAMIC}\text{-}\mathsf{Else})$$

$$\mathsf{t} \rightarrow \mathsf{t}' \qquad (\mathsf{E}\text{-}\mathsf{DYNAMIC})$$

Progress: Suppose t is a well-typed term (that is, t:T for some T). Then either t is a value or else there is some t' with $t \to t'$. Missing cases:

- Case (T-Dynamic): t = dynamic t' as T, t': T
 By the induction hypothesis, either t' is a value or else there is some t" such that t' → t".
 If t' is a value then the whole packaged dynamic value dynamic t' as T is a value. Otherwise rule E-Dynamic applies.
- Case (T-Typecase): $t = \text{typecase } t' \text{ match } \{ \text{ case } x_i : T_i \Rightarrow t_i \text{ else } t_{else} \}$ By the induction hypothesis, either t' is a value or else there is some t'' such that $t' \to t''$. If t' is not a value then rule (E-Typecase) applies. If t' is a value and from premises we have that t': T' then by canonical forms lemma we know that t' must be a dynamic value of a form dynamic t'' as T'. Then either there exists a tag T_i which matches T' and rule (E-Typecase-Dynamic) applies, or there is none. For the latter case we have a fallback else case, and rule (E-Typecase-Dynamic-Else) applies.

Preservation (for completeness): If $\Gamma \vdash t : T$ and $t \to t'$, then $\Gamma \vdash t' : T$ Missing cases:

- Case (T-Dynamic): t = dynamic t₁ as T
 From the premise of (T-Dynamic) we have that Γ ⊢ t₁: T. The only evaluation rule which applies is (E-Dynamic). And then t' is of a form t' = dynamic t'₁ as T and by induction hypothesis t'₁ is of type T. By (T-Dynamic) Γ ⊢ dynamic t'₁ as T: T.
- Case (T-Typecase): $t = \text{typecase } t_1 \text{ match } \{ \text{case } x_i : T_i \Rightarrow t_i \text{ else } t_{else} \}$ From the premise of (T-Typecase) we have that $t_1 : \text{Dynamic}$. Looking at the evaluation rules with typecase on the LHS we have three cases to consider:
 - Case (E-Typecase): A reduced term is of the form $\mathbf{t}' = \mathbf{typecase} \ \mathbf{t}'_1 \ \mathsf{match} \ \{ \ \mathsf{case} \ \mathbf{x}_i : \mathsf{T}_i \Rightarrow \mathbf{t}_i \ \mathsf{else} \ \mathsf{t}_{else} \}$ and by induction hypothesis we know that $\mathbf{t}'_1 : \mathsf{Dynamic}$. Since other premises (for cases) remain the same we have that $\mathbf{t}' : \mathsf{T}$ by (T-Typecase).
 - Case (E-Typecase-Dynamic): By the premise of the evaluation rule we have that $t_1 = \text{dynamic } v$ as T_v and the typing rule $\Gamma, x_k : T_k \vdash t_k : T$. The RHS of the evaluation rule gives us $t' = [x_k \to v] t_k$. By the premise of k'th case we know the type tag of the dynamic value $T_v = T_k$. Therefore by the substitution lemma we have that $\Gamma, x_k : T_v \vdash [x_k \to v] t_k : T$
 - Case (E-Typecase-Dynamic-Else): By the premise of the rule we know that none of the cases matched. Hence else case applies with $\Gamma \vdash \mathbf{t}_{else}$: T by the premise of (T-Typecase).

q.e.d

For reference: predefined lambda terms

Predefined lambda terms that can be used as-is in "Hacking with the untyped call-by-value lambda calculus":

```
\lambda t. \lambda f. t
tru =
fls =
                  \lambda t. \ \lambda f. \ f
iszro = \lambda m. m (\lambda x. fls) tru
                  \lambda b. \ \lambda t. \ \lambda f. \ b \ t \ f \ unit
\mathsf{test} =
                  \lambda f. \ \lambda s. \ \lambda b. \ b \ f \ s
pair =
fst =
                  \lambda p. p tru
\mathtt{snd} =
                  \lambda p. p fls
                  \lambda s. \ \lambda z. \ z
c_0 =
                  \lambda s. \ \lambda z. \ s \ z
c_1 =
scc =
                  \lambda n. \ \lambda s. \ \lambda z. \ s \ (n \ s \ z)
                  \lambda m. \ \lambda n. \ \lambda s. \ \lambda z. \ m \ s \ (n \ s \ z)
plus =
times = \lambda m. \lambda n. m (plus n) c_0
zz =
                  pair c_0 c_0
                  \lambda p. \ pair \ (snd \ p) \ (scc \ (snd \ p))
ss =
                  \lambda m. \ fst \ (m \ ss \ zz)
prd =
                  \lambda f. (\lambda x. f (\lambda y. x x y)) (\lambda x. f (\lambda y. x x y))
fix =
```

The call-by-value simply typed lambda calculus

The complete reference of the variant of simply typed lambda calculus (with Bool ground type representing the type of values true and false) used in "The call-by-value simply typed lambda calculus with returns" and "Dynamic" is as follows:

$$\begin{array}{lll} v ::= \lambda x : T. \ t \mid bv & (values) \\ bv ::= \mathtt{true} \mid \mathtt{false} & (boolean \ values) \\ t ::= x \mid v \mid t \ t & (terms) \\ p ::= t & (programs) \\ T ::= \mathtt{Bool} \mid T \rightarrow T & (types) \end{array}$$

Evaluation rules:

$$\frac{t_1 \longrightarrow t_1'}{t_1 \ t_2 \longrightarrow t_1' \ t_2} \tag{E-APP1}$$

$$\frac{t_2 \longrightarrow t_2'}{v_1 \ t_2 \longrightarrow v_1 \ t_2'} \tag{E-App2}$$

$$(\lambda x \colon \mathsf{T}_1 \colon t_1) \ v_2 \longrightarrow [x \to v_2]t_1$$
 (E-Appabs)

Typing rules:

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

$$\frac{\Gamma, \ x : \mathsf{T}_1 \vdash t_2 : \mathsf{T}_2}{\Gamma \vdash (\lambda x \colon \mathsf{T}_1 \cdot t_2) : \mathsf{T}_1 \to \mathsf{T}_2} \tag{T-Abs}$$

$$\frac{\Gamma \vdash t_1 : \mathsf{T}_1 \to \mathsf{T}_2 \quad \Gamma \vdash t_2 : \mathsf{T}_1}{\Gamma \vdash t_1 \ t_2 : \mathsf{T}_2} \tag{T-App}$$

$$\frac{}{\Gamma \vdash \mathsf{true} : \mathsf{Bool}} \tag{T-False}$$

$$\Gamma \vdash \mathtt{false} : \mathtt{Bool}$$
 (T-True)