

# 1 Checked Error Handling

In this exercise we use the Simply-Typed Lambda Calculus (STLC) extended with rules for error handling. In this language, terms may reduce to a normal form **error**, which is *not* a value. In addition, we add the new term form **try**  $t_1$  **with**  $t_2$ , which allows handling errors that occur while evaluating  $t_1$ .

Here is a summary of the extensions to syntax and evaluation:

$t ::=$	<b>terms :</b>
...	
<b>error</b>	run-time error
<b>try</b> $t$ <b>with</b> $t$	trap errors

New evaluation rules:

$$\begin{array}{ll}
 \mathbf{error} \ t_2 \longrightarrow \mathbf{error} \text{ (E-APPERR1)} & \mathbf{try} \ v_1 \ \mathbf{with} \ t_2 \longrightarrow v_1 \text{ (E-TRYVALUE)} \\
 v_1 \ \mathbf{error} \longrightarrow \mathbf{error} \text{ (E-APPERR2)} & \mathbf{try} \ \mathbf{error} \ \mathbf{with} \ t_2 \longrightarrow t_2 \text{ (E-TRYERROR)} \\
 \frac{t_1 \longrightarrow t'_1}{\mathbf{try} \ t_1 \ \mathbf{with} \ t_2 \longrightarrow \mathbf{try} \ t'_1 \ \mathbf{with} \ t_2} & \text{(E-TRY)}
 \end{array}$$

These extensions are exactly those summarized in Figures 14-1 and 14-2 on pages 172 and 174 of the TAPL book. However, we will use *different* typing rules, because we want a stronger progress result.

Unlike the relaxed progress statement used by TAPL, we want to ensure that all errors are eventually caught, i.e., that no error reaches the top-level. For example, we want to accept **try** **error** **with** **true** but we want to reject  $(\lambda x:\mathbf{Bool}.x) \ \mathbf{error}$ .

The goal of this exercise is to prove that the following progress theorem holds:

**Theorem 1** (Progress). *If  $\Gamma ; \mathbf{false} \vdash t:T$ , then either  $t$  is a value or else  $t \longrightarrow t'$ .*

The above theorem uses a typing judgment extended with a Boolean value  $E$ , written

$$\Gamma ; E \vdash t:T$$

where  $E \in \{\mathbf{true}, \mathbf{false}\}$ .

The theorem says that a well-typed term that is closed (that is, it does not have free variables, which is expressed using  $\Gamma = \emptyset$ ) is either a value, or else it can be reduced *as long as*  $E = \mathbf{false}$ .

Intuitively, the predicate  $E$  in the typing judgment  $\Gamma ; E \vdash t:T$  determines whether the term  $t$  is in an *impure* position, i.e. whether it is allowed to reduce to **error** or not.

The typing rules inherited from STLC “push” the impurity predicate  $E$  unchanged from the conclusion into the premises. Intuitively, if a term is allowed to produce an error, all its sub-terms are also allowed to do so.

$$\begin{array}{ll}
 \frac{x:T \in \Gamma}{\Gamma ; E \vdash x:T} & \text{(T-VAR)} \\
 \frac{\Gamma, x:T_1 ; E \vdash t_2:T_2}{\Gamma ; E \vdash \lambda x:T_1.t_2:T_1 \rightarrow T_2} & \text{(T-ABS)} \\
 \frac{\Gamma ; E \vdash t_1:T_1 \rightarrow T_2 \quad \Gamma ; E \vdash t_2:T_1}{\Gamma ; E \vdash t_1 \ t_2:T_2} & \text{(T-APP)}
 \end{array}$$

Your task is as follows:

1. Design typing rules for **error** (T-ERROR) and **try t with t** (T-TRY) terms, so that you are able to prove the progress theorem (and so that it is possible to write reasonable programs).
2. Prove the above progress theorem using structural induction. (You can use the canonical forms lemma for STLC as seen in the lecture without proof.)
3. How does  $E$  contribute to your proof? Where would your proof “go wrong” if we removed it from the typing rules?
4. (*Extra effort*) Try to prove the preservation theorem as well. Is it possible with the typing rules you designed? If not, is there a way to alter the typing rules to make the theorem hold? In either case, it may be useful to slightly change the form of the function type.

## 1.1 Solution

### 1.1.1 Typing rules

$$\Gamma ; \text{true} \vdash \text{error} : T \quad (\text{T-ERROR})$$

$$\frac{\Gamma ; \text{true} \vdash t_1 : T \quad \Gamma ; E \vdash t_2 : T}{\Gamma ; E \vdash \text{try } t_1 \text{ with } t_2 : T} \quad (\text{T-TRY})$$

### 1.1.2 Progress of pure terms.

We have to prove: if  $\Gamma ; \text{false} \vdash t : T$ , then either  $t$  is a value or else  $t \longrightarrow t'$ .

*Proof.* The proof is by induction on typing derivations and resembles again that of the progress theorem for the STLC (see TAPL, Thm. 9.3.5, p. 105).

CASE T-VAR, T-ABS, T-APP

Identical to the proof for STLC, modulo the appearance of  $E = \text{false}$  in the context.

CASE T-ERROR

Cannot happen: T-ERROR requires  $E = \text{true}$ .

CASE T-TRY

We know:  $t = \text{try } t_1 \text{ with } t_2$ ,  $\Gamma ; \text{true} \vdash t_1 : T$  and  $\Gamma ; \text{false} \vdash t_2 : T$ .

The induction hypothesis for  $t_1$  would be: if  $\Gamma ; \text{false} \vdash t_1 : T_1$ , then either  $t_1$  is a value or else  $t_1 \longrightarrow t'_1$ . Unfortunately, we do not have  $\Gamma ; \text{false} \vdash t_1 : T$ , so we cannot use it!

Clearly, we must allow  $t_1$  to evaluate to an error here, in addition to being a value or taking a step. This observation leads to the following “weak progress” lemma:

**Lemma 1** (Weak Progress). *If  $\Gamma ; E \vdash t : T$ , then either (a)  $t$  is a value, (b)  $t = \text{error}$ , or (c)  $t \longrightarrow t'$ .*

Using Weak Progress on  $t_1$ , we get 3 subcases:

- If  $t_1$  is a value, E-TRYVALUE applies.
- If  $t_1 = \text{error}$ , E-TRYERROR applies.
- If  $t_1 \longrightarrow t'_1$ , E-TRY applies.

□

We now have to prove the Weak Progress lemma.

*Proof.* The proof is similar to a progress proof, and works by induction on the type derivation.

CASE T-VAR, T-ABS, T-APP

Identical to the proof for STLC, modulo the appearance of  $E$  in the context.

CASE T-ERROR

Immediate:  $t = \text{error}$ .

CASE T-TRY

We know:  $t = \text{try } t_1 \text{ with } t_2, \Gamma ; \text{true} \vdash t_1 : T$  and  $\Gamma ; E \vdash t_2 : T$ .

We now have an induction hypothesis on  $t_1$  that we can use, and which tells us that either  $t_1$  is a value,  $t_1 = \text{error}$ , or  $t_1 \longrightarrow t'_1$ .

- If  $t_1$  is a value, E-TRYVALUE applies.
- If  $t_1 = \text{error}$ , E-TRYERROR applies.
- If  $t_1 \longrightarrow t'_1$ , E-TRY applies.

□

## 2 The call-by-value simply typed lambda calculus with returns

Consider a variant of the call-by-value simply typed lambda calculus 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.

This system extends STLC with booleans with two new forms of terms:

$$\begin{array}{lcl} t & ::= & \dots \quad \text{terms :} \\ & | & \text{return } t \quad \text{return} \\ & | & \text{frame } t \quad \text{frame} \end{array}$$

**frame** is not meant to be used in terms representing source programs. Rather, it is introduced with the following reduction rule:

$$(\lambda x:T_1. t_{12}) \ v \longrightarrow \text{frame } ([x \mapsto v] t_{12}) \quad (\text{E-APPABS})$$

The frame form marks the scope from which local returns should return. If the body of a frame reduces to a value, the frame should be removed. Otherwise, a return form should stop propagating at a frame in much the same way that an error stops at a **try/with**.

In the previous exercise, we extended the typing rules to track impurity. We do something similar here, but to track the types of values that can be *returned* from a term. That is in addition to the normal type of the term, which is the one we get if it finishes evaluating without encountering a **return**. We use  $R$  to denote a *set* of types, i.e.,  $\{T_1, \dots, T_n\}$ .

$$\begin{array}{c} \Gamma \vdash \text{true} : \text{Bool} \mid \emptyset \quad (\text{T-TRUE}) \qquad \frac{\Gamma \vdash t_1 : T_1 \rightarrow T_2 \mid R_1 \quad \Gamma \vdash t_2 : T_1 \mid R_2}{\Gamma \vdash t_1 \ t_2 : T_2 \mid (R_1 \cup R_2)} \quad (\text{T-APP}) \\ \Gamma \vdash \text{false} : \text{Bool} \mid \emptyset \quad (\text{T-FALSE}) \\ \\ \frac{x : T \in \Gamma}{\Gamma \vdash x : T \mid \emptyset} \quad (\text{T-VAR}) \qquad \frac{\Gamma \vdash t_1 : T_1 \mid R_1}{\Gamma \vdash \text{return } t_1 : T \mid (R_1 \cup \{T_1\})} \quad (\text{T-RETURN}) \\ \\ \frac{\Gamma, x : T_1 \vdash t_2 : T_2 \mid R_2 \quad R_2 \subseteq \{T_2\}}{\Gamma \vdash \lambda x : T_1. t_2 : T_1 \rightarrow T_2 \mid \emptyset} \quad (\text{T-ABS}) \qquad \frac{\Gamma \vdash t_1 : T \mid R_1 \quad R_1 \subseteq \{T\}}{\Gamma \vdash \text{frame } t_1 : T \mid \emptyset} \quad (\text{T-FRAME}) \end{array}$$

Note that the conclusion of T-RETURN mentions a  $T$  that does not appear anywhere in the premises. What does that mean?

In this exercise, you are to given typing rules for abstractions and frames, and adjust the existing evaluation rules such that they correctly and comprehensively describe the semantics of the extension. In particular, **return** terms must short-circuit most of the terms, much like **error** terms did in the previous exercise.

You must then prove the preservation theorem for the system.

More precisely, you have the following tasks:

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. Make sure that **return** terms propagate like **error** terms did in the previous exercise, so that progress would hold.
2. Design typing rules for abstractions (T-ABS) and frames (T-FRAME) so that you are able to prove the following lemmas and preservation theorem (and so that it is possible to write reasonable programs).

3. Prove the following two lemmas:

Values are pure: If  $\Gamma \vdash v : T \mid R$ , then  $R = \emptyset$ .

**return** is polymorphic: If  $\Gamma \vdash \text{return } t_1 : T \mid R$ , then  $\Gamma \vdash \text{return } t_1 : U \mid R$ .

4. State, but do not prove, the weakening lemma, the permutation lemma, and the substitution lemma for this system.
5. State and prove the preservation theorem. You may rely on the above lemmas as well as the usual inversion lemma and canonical forms lemma. Think about how  $R$  evolves when we take one step of evaluation: does it always remain the same?
6. (*Extra effort*) State and prove the progress theorem as well. The statement should exclude top-level returns – intuitively, a closed term with a “naked” return doesn’t make sense.

Before you begin, think carefully about the following simple term:  $\lambda x:\text{Bool}.(\text{return true}) \ x$ . Intuitively, it makes sense. Once this lambda is applied, it is going to evaluate to **true**, regardless of the input. What type or types will be assigned to **return true**?

## 2.1 Solution

### 2.1.1 Evaluation rules

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

$$\frac{t_2 \longrightarrow t'_2}{v_1 \ t_2 \longrightarrow v_1 \ t'_2} \quad (\text{E-APP2})$$

$$(\lambda x : T_1 . t_2) \ v_1 \longrightarrow \text{frame } ([x \mapsto v_1]t_2) \quad (\text{E-APPABS})$$

$$\frac{t_1 \longrightarrow t'_1}{\text{return } t_1 \longrightarrow \text{return } t'_1} \quad (\text{E-RET})$$

$$(\text{return } v_1) \ t_2 \longrightarrow \text{return } v_1 \quad (\text{E-APPRET1})$$

$$v_1 \ (\text{return } v_2) \longrightarrow \text{return } v_2 \quad (\text{E-APPRET2})$$

$$\text{return } (\text{return } v_1) \longrightarrow \text{return } v_1 \quad (\text{E-RETRRET})$$

$$\frac{t_1 \longrightarrow t'_1}{\text{frame } t_1 \longrightarrow \text{frame } t'_1} \quad (\text{E-FRAME})$$

$$\text{frame } v_1 \longrightarrow v_1 \quad (\text{E-FRAMEVAL})$$

$$\text{frame } (\text{return } v_1) \longrightarrow v_1 \quad (\text{E-FRAMERET})$$

### 2.1.2 Values are pure

Values-are-pure lemma: If  $\Gamma \vdash v : T \mid R$ , then  $R = \emptyset$ .

Proof: immediate from T-TRUE, T-FALSE and T-ABS.

### 2.1.3 return is polymorphic

Return-is-poly lemma: If  $\Gamma \vdash \mathbf{return} \ t_1 : T \mid R$ , then  $\Gamma \vdash \mathbf{return} \ t_1 : U \mid R$ .

Proof: from T-RETURN, we know there must exist  $T_1$  and  $R_1$  such that  $\Gamma \vdash t_1 : T_1 \mid R_1$  with  $R = R_1 \cup \{T_1\}$ . Then, by T-RETURN, we have  $\Gamma \vdash \mathbf{return} \ t_1 : U \mid R_1 \cup \{T_1\}$ .

### 2.1.4 Additional lemmas

Weakening lemma: If  $\Gamma \vdash t : T \mid R$  and  $x \notin \text{dom}(\Gamma)$ , then  $\Gamma, x : T_1 \vdash t : T \mid R$ .

Permutation lemma: If  $\Gamma \vdash t : T \mid R$  and  $\Delta$  is a permutation of  $\Gamma$ , then  $\Delta \vdash t : T \mid R$ .

Substitution lemma: If  $\Gamma, x : S \vdash t : T \mid R$  and  $\Gamma \vdash s : S \mid \emptyset$ , then  $\Gamma \vdash [x \mapsto s]t : T \mid R$ .

### 2.1.5 Proof of preservation

Preservation theorem:

If  $\Gamma \vdash t : T \mid R$  and  $t \longrightarrow t'$ , then  $\Gamma \vdash t' : T \mid R'$  with  $R' \subseteq R$ .

Proof: by induction on the typing derivation tree. We proceed by case analysis of the last typing rule in the derivation tree.

CASE T-TRUE, T-FALSE, T-VAR, T-ABS and T-RETURN :

Trivial, because they are all for normal forms, i.e., there is no  $t'$  such that  $t \longrightarrow t'$ .

CASE T-FRAME :

We know:  $t = \mathbf{frame} \ t_1$ ,  $\Gamma \vdash t_1 : T \mid R_1$ ,  $R_1 \subseteq \{T\}$ ,  $R = \emptyset$ .

We proceed by case analysis for the derivation of  $t \longrightarrow t'$ .

SUBCASE E-FRAME :

We know:  $t_1 \longrightarrow t'_1$ ,  $t' = \mathbf{frame} \ t'_1$ .

By the induction hypothesis, we have  $\Gamma \vdash t'_1 : T \mid R'_1$  with  $R'_1 \subseteq R_1$ . Then by T-FRAME and because  $R'_1 \subseteq R_1 \subseteq \{T\}$ , we have  $\Gamma \vdash \mathbf{frame} \ t'_1 : T \mid \emptyset$ , as required.

SUBCASE E-FRAMEVAL :

We know:  $t_1 = v_1$ ,  $t' = v_1$ .

By the values-are-pure lemma, we know that  $R_1 = \emptyset \subseteq R$ . This combines with  $\Gamma \vdash t_1 : T \mid R_1$  to conclude.

SUBCASE E-FRAMERET :

We know:  $t_1 = \mathbf{return} \ v_1$ ,  $t' = v_1$ .

From T-RETURN, we know that there exist  $T_1$  and  $Q$  such that  $\Gamma \vdash v_1 : T_1 \mid Q$  and  $R_1 = Q \cup \{T_1\}$ . By the values-are-pure lemma, we know that  $Q = \emptyset$ , hence  $R_1 = \{T_1\}$ . Since  $R_1 \subseteq \{T\}$ , we have  $\{T_1\} \subseteq \{T\}$  and hence  $T_1 = T$ . This allows us to conclude that  $\Gamma \vdash v_1 : T \mid \emptyset$ , as required.

CASE T-RETURN :

We know:  $t = \mathbf{return} \ t_1$ ,  $\Gamma \vdash t_1 : T_1 \mid R_1$ ,  $R = R_1 \cup \{T_1\}$ .

We proceed by case analysis for the derivation of  $t \longrightarrow t'$ .

SUBCASE E-RET :

We know:  $t_1 \longrightarrow t'_1$ ,  $t' = \mathbf{return} \ t'_1$ .

By the induction hypothesis, we have  $\Gamma \vdash t'_1 : T_1 \mid R'_1$  with  $R'_1 \subseteq R_1$ . Then by T-RETURN, we have  $\Gamma \vdash \text{return } t'_1 : T_1 \mid (R'_1 \cup \{T_1\})$ . Because  $R'_1 \subseteq R_1$ , we have that  $R'_1 \cup \{T_1\} \subseteq R$ , which concludes.

SUBCASE E-RETRET :

We know:  $t_1 = \text{return } v_{11}, t' = \text{return } v_{11}$ .

By the polymorphic-return lemma, we have  $\Gamma \vdash t_1 : T \mid R_1$ . We conclude by noting that  $t' = t_1$  and  $R_1 \subseteq R$ .

CASE T-APP :

We know:  $t = t_1 t_2, \Gamma \vdash t_1 : T_1 \rightarrow T \mid R_1, \Gamma \vdash t_2 : T_1 \mid R_2, R = R_1 \cup R_2$ .

We proceed by case analysis for the derivation of  $t \rightarrow t'$ .

SUBCASE E-APP1 :

We know:  $t_1 \rightarrow t'_1, t' = t'_1 t_2$ .

By the induction hypothesis, we have  $\Gamma \vdash t'_1 : T_1 \mid R'_1$  with  $R'_1 \subseteq R_1$ . Then by T-APP, we have  $\Gamma \vdash t'_1 t_2 : T \mid (R'_1 \cup R_2)$ . Since  $R'_1 \subseteq R_1$ , we have that  $(R'_1 \cup R_2) \subseteq R$ , which concludes.

SUBCASE E-APP2 :

Similar.

SUBCASE E-APPRET1 :

We know:  $t_1 = \text{return } v_1, t' = \text{return } v_1$ .

By the polymorphic-return lemma, we have  $\Gamma \vdash t_1 : T \mid R_1$ . We conclude by noting that  $t' = t_1$  and  $R_1 \subseteq R$ .

SUBCASE E-APPRET2 :

Similar.

SUBCASE E-APPABS :

We know:  $t_1 = \lambda x : T_{11}. t_{12}, t_2 = v_2, t' = \text{frame } [x \mapsto v_2]t_{12}$ .

From T-ABS,  $T_{11} = T_1$  and  $\Gamma, x : T_1 \vdash t_{12} : T \mid Q$  with  $Q \subseteq \{T\}$ . By the values-are-pure lemma,  $R_2 = \emptyset$ , i.e.,  $\Gamma \vdash v_2 : T_1 \mid \emptyset$ . Therefore, by the substitution lemma, we have  $\Gamma \vdash [x \mapsto v_2]t_{12} : T \mid Q$ . Combining with  $Q \subseteq \{T\}$ , we use T-FRAME to conclude that  $\Gamma \vdash \text{frame } [x \mapsto v_2]t_{12} : T \mid \emptyset$ , as required.

### 2.1.6 Commentary

The particular choice of typing rules we used is not the only possible solution. The typing rules presented treat  $R$  like their output, something that we can determine from the environment, term and the premise.

In (T-VAR), the return set is always empty, because there's nothing else we can reasonably put there. It makes sense when applying the substitution in E-APPABS, since we replace variables by values, which are pure. In (T-RETURN), we *extend* the current returns with the type of the value we're currently trying to return. In (T-APP), we just take the union of the return sets of the two terms we have. In T-ABS, where we must actually “merge” a set  $R_2$  with a single type  $T_2$ , we use the slight trick of demanding that  $R_2 \subseteq \{T_2\}$ , which is equivalent to demanding that  $R_2$  be empty or equal to the singleton  $\{T_2\}$ .

So do we need  $R$  to be a set in this task? We want  $R$  to sometimes be empty (because some expressions never explicitly **return**), and we want a convenient notion of “extending”  $R$  with another type. A set works quite well here. Try to think how you'd reformulate rules such as (T-APP) without using sets.

### 3 Appendix

#### 3.1 The call-by-value simply typed lambda calculus with booleans

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” is as follows:

<b>t ::=</b>	<b>terms :</b>
x	variable
$\lambda x:T.t$	abstraction
t t	application
true   false	boolean values
if t then t else t	conditional
<b>v ::=</b>	<b>values :</b>
$\lambda x:T.t$	abstraction
true   false	boolean values
<b>T ::=</b>	<b>types :</b>
$T \rightarrow T$	function type
Bool	boolean type

Evaluation rules:

$\frac{t_1 \longrightarrow t'_1}{t_1 \ t_2 \longrightarrow t'_1 \ t_2} \quad (\text{E-APP1})$	$\text{if true then } t_2 \text{ else } t_3 \longrightarrow t_2 \quad (\text{E-IFTRUE})$
$\frac{t_2 \longrightarrow t'_2}{v_1 \ t_2 \longrightarrow v_1 \ t'_2} \quad (\text{E-APP2})$	
$(\lambda x:T.t_{12}) \ v_2 \longrightarrow [x \mapsto v_2]t_{12} \quad (\text{E-APPABS})$	$\text{if true then } t_2 \text{ else } t_3 \longrightarrow t_3 \quad (\text{E-IFFALSE})$
$\frac{t_1 \longrightarrow t'_1}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \longrightarrow \text{if } t'_1 \text{ then } t_2 \text{ else } t_3} \quad (\text{E-IF})$	

Typing rules:

$\frac{x:T \in \Gamma}{\Gamma \vdash x:T} \quad (\text{T-VAR})$	$\Gamma \vdash \text{true}: \text{Bool} \quad (\text{T-TRUE})$
$\frac{\Gamma, x:T_1 \vdash t_2:T_2}{\Gamma \vdash (\lambda x:T_1.t_2):T_1 \rightarrow T_2} \quad (\text{T-ABS})$	$\Gamma \vdash \text{false}: \text{Bool} \quad (\text{T-FALSE})$
$\frac{\Gamma \vdash t_1:T_1 \rightarrow T_2 \quad \Gamma \vdash t_2:T_1}{\Gamma \vdash t_1 \ t_2:T_2} \quad (\text{T-APP})$	$\frac{\Gamma \vdash t_1:\text{Bool} \quad \Gamma \vdash t_2:T \quad \Gamma \vdash t_3:T}{\Gamma \vdash \text{if } t_1 \text{ then } t_2 \text{ else } t_3:T} \quad (\text{T-IF})$