Verifying correctness of Stainless programs using Coq

Bence Czipó, Jad Hamza

June 8, 2018

1 Introduction

Correctness of programs is hard to verify

2 Background

2.1 Stainless

Stainless is a verification framework for Scala programs. Among other things, it extends scala functions with the notion of pre- and postcondition. From the practical approach, a Stainless program can be divided into two parts: ADT definitions and functions. A general stainless function consists of a precondition pre, a function body body and a postcondition λres. post(res) where

$$\text{pre} \implies (\lambda\text{res. post(res)})\text{body}$$

2.2 Coq

Coq is an proof assistant based on the calculus of constructions. It provides functionality to write definitions and theorems and an interactive environment to prove them. Even though Coq is not an automated theorem prover, it also provides some automatism through tactics.

One big advantage of Coq is that there exist several libraries that extend the core features. During our work, we used two of those.

2.2.1 Program Library

Program is a Coq library that introduces the concept of dependent types in a form of $\{e : T | P\}$ where T is a type and P is a predicate.

The program library generates obligations to plug in the "holes" in the types. The environment tries to solve the generated obligation with a user-defined obligation tactic.

2.2.2 Coq-Equations

"Equations provides a notation for writing programs by dependent pattern-matching and (well-founded) recursion in Coq. It compiles everything down to

eliminators for inductive types, equality and accessibility, providing a definitional extension to the Coq kernel."

2.2.3 Coq-std++

Coq-std++ (stdpp) is a self-contained collection of data structures, lemmas and tactics that is intended to extend the functionality of the standard library of Coq. It fetaures a set representation with all the common operations and a solver tactic to solve goals involving them.

3 Transforming Stainless Programs to Coq

In order to verify Stainless programs in Coq, they are needed to be transformed into a correctly typed Coq program.

Definition 1 (Translation function) *Let t be a function that assigns a (correctly typed) Coq program to every (correctly typed) Stainless program. Though t is not designed to be invertible, for simplicity let us introduce the t^{-1} notation where $t^{-1}(t(p)) = p$. Also for simplicity, let us denote $t(p)$ by $[p]$.*

Many scala construct has its exact representation in Coq. For instance integers can be handled with the Z library, lambda constructs can be translated to coq lambdas, or let's can be translated directly into lets in coq. In the following we will only highlight the non-trivial translation steps. Also, the uniqueness of function and parameter names has to be ensured that we achieve through renaming. From now on, let us assume that every name is unique in its scope. The if-then-else expressions in Scala could be translated into Coq's own if-then-else constructs, however, during verification, we would like to propagate the boolean value of the condition into the branches. For this reason, we defined our own if-then-else constructs.

```
Definition ifthenelse b A (e1: true = b → A) (e2: false = b → A): A :=
match b as B return (B = b → A) with
| true ⇒ fun H ⇒ e1 H
| false ⇒ fun H ⇒ e2 H
end eq_refl.
```

We use this if-then-else to translate non-exhaustive matches and boolean and operations. Other boolean operations are translated to their Coq correspondent. The sets and set operations are all translated to the structures of the stdpp library. Coq equality is translated to the = operator in Coq with exception of booleans, integers and sets. The equality operator has Prop result, however, the context in which we use equality usually requires booleans. To convert the two, we defined a propInBool function, however it is also requires to assume a classical type system, i.e. every Prop is either True or False.

3.1 Translating ADT's

The translation of whole programs consists of two steps: translating ADT's and translating functions.

In stainless, there are two kinds of ADT's, ADTSort for the root of class hierarchies and ADTConstructor for the case classes.

For example consider the following class hierarchy:

```scala
sealed abstract class List[T] {}

case class Nil[T]() extends List[T] {}

case class Cons[T](h: T, t: List[T]) extends List[T] {}
```

The translation starts from each ADTSort and creates an inductive definition stating that an ADTSort is one of its constructor. A constructor can be expressed as a type, that takes the types of arguments and maps it to an ADTSort. In our example, the translated version would be.

```coq
Inductive List (T: Type) :=
| Cons: T → ((List T) → (List T))
| Nil: List T.
```

Based on the constructor used to construct the object, we can define a recognizer, that decides for an abstract supertype if it is instance of a concrete subtype. This can be achieved through pattern matching. For example in case of Cons it would be:

```coq
Definition isCons (T: Type) (src: List T) : bool :=
match src with
| Cons _ _ _ ⇒ true
| _ ⇒ false
end.
```

Using this recognizers, we can also define a type for each subtype as a dependent type. For Cons from our example, it would be:

```coq
Definition Cons_type (T: Type) : Type :=
{src: List T | (isCons T src = true)}.
```

Now, with types corresponding to case classes, we can express the field accessors by simply pattern matching on the object. The corresponding argument of the constructor can be reutrned. For instnace in case of the ead of a list it would be:

```coq
Definition h (T: Type) (src: Cons_type T) : T :=
match src with
| Cons_construct _ f0 f1 ⇒ f0
| _ ⇒ let contradiction: False := _ in match contradiction with end
end.
```

Note that the second branch is only required so that the match is exhaustive. However due to obligations, that branch is impossible, and the environment is forced to check that with deriving False.

We also generate some lemmas and tactics using them to rewrite certain expressions using ADT constructs that are not detailed here.

3.2 Translating Functions

Since forward declaration is difficult to implement in Coq, it is important to have a total order of functions, in which no function is referenced before it is defined. For this reason, our transforming methods does not support mutually recursive functions. It is easy to see that if there is no mutual recursivity in a set of functions, such order exists.

3.2.1 Translating Non-recursive Functions

A function f with body b, type parameters $T_1, \ldots T_k$, parameters $p_1, \ldots p_n$ with types $U_1, U_2, \ldots U_n$ and a return type U_r can be translated into a simple Coq definition. Preconditions can be interpreted as assertion in the beginning of the function. This can be expressed in Coq by taking an argument that states that the precondition holds, in oder means, taking an argument with the type pre = true. Postconditions can be expressed by a dependent return type. In case of a postcondition λres. post(res) the return type changes to $\{res : U_r \mid post(res)\}$. So a general translated function looks like this:

```
Definition f (T₁: Type) ... (Tₖ: Type)(p₁: [U₁]) ... (pₙ: [Uₙ]) (prec1:
    [pre] = true) :
    {res: [Uᵣ] | [post] res} :=
        [b].
```

3.2.2 Translating Recursive Functions

The Program library offers `Program Fixpoint` to handle recursive functions, however, the fixpoint operator makes it difficult to rewrite recursive functions with their definitions. To overcome this, we used the Equations library.

In case of recursive functions, we decided to focus on verification, so we ignored termination. For this, we defined an axiom, that we can pass as a decreasing measure:

```
    Axiom ignore_termination: nat.
```

Equations do not allow dependent types like the Program library did. In order to work around that, in case of a postcondition, we have to define the return type using a Program Definition, and after, we can use that type in the equations. Thus in this case, a translated function would be:

```
Definition rt (T₁: Type) ... (Tₖ: Type)(p₁: [U₁]) ... (pₙ: [Uₙ]) (prec1:
    [pre] = true) ... : Type :=
    {res: [Uᵣ] | [post] res}

Equations f (T₁: Type) ... (Tₖ: Type)(p₁: [U₁]) ... (pₙ: [Uₙ]) (prec1:
    [pre] = true) ... :
  rt T₁ ... Tₖ p₁ ... pₙ prec1 ... :=
    f T₁ ... Tₖ p₁ ... pₙ prec1 ... by rec ignore_termination lt :=
    f T₁ ... Tₖ p₁ ... pₙ prec1 ... :=
        [b].
```

Later, if we want to rewrite with the content of the function, we can rely on the fact that it will be expressed by `f_equation_1`.

3.2.3 Function Invocation

The function application can be translated as function application in Coq, however, it must be taken into account, that functions with preconditions take more parameters. The program library allows us to pass unknown arguments (underscores) and requires us to construct proof for them later in form of obligation proofs. When the invoked function has a postcondition the return type is not exactly what is a refined type, that needs to be projected.

4 Correctness Equivalence

We only checked two aspects of validity of stainless programs, so let us define validity accordingly

Definition 2 (Stainless-validity) *A Stainless program p is valid if*

- *the preconditions hold for every call of f*

- *for any input satisfying the preconditions, the postcondition holds*

In stainless, pre- and postconditions can be expressed as evaluating a boolean condition, and if the condition does not hold, go into an error branch. So in other terms, p is correct, if these error branches are never active.

Errors are translated to Coq using contradiction: [error] = contradiction _
Contradictions can be expressed as the obligation to generate false.

```
Definition contradiction(p: False) : T := match p with end.
```

Applying `contradiction` on an unknown variable will result in an obligation to derive `False` from the context.

Definition 3 (Coq-validity) *A Coq program p is valid if there is a proof (an expression of that type) for every obligation generated by it.*

The goal of the translation presented in Section 3 was to be able to reason about the validity (or correctness) of the original Stainless program, based on the validity of the translated Coq representation.

Theorem 1 *For every (correctly typed) Stainless program p and its translation [p], if [p] is proved to be valid by Coq, than p is also valid.*

Instead of proving this we will just sketch some notions why is it true.

Assumption 1 *For every every Scala term `b: Boolean` we assume that if $b \rightarrow^* c$ where $c \in \{true, false\}$ then $[b] \rightarrow^*_{coq} [c]$, so if b evaluates to a boolean constant in Stainless (under the context Γ), its translated representation evaluate to the representation of that boolean constant in Coq (under the context $[\Gamma]$).*

Let us examine only the cases where every decision in the program p is a branching based on a boolean variable. We will also assume that p does not have any free variables. Our statement can be proved by contradiction, assuming that $[p]$ is proven to be correct, but p crashes. In this case, there is a branch in p that contains error, and it is accessed through evaluating the boolean expressions $b_1, b_2 \ldots b_n$. We can apply Assumption 1, add conclude that $[p]$ reduces to the same branch. We know that Coq was able to verify $[p]$, which means it was able to prove the obligation that was generated by contradiction. This means that we had $\{\} \vdash x : False$ as hypothesis.

Here, we can use two facts:

Fact 1 *In Coq, if $\{\} \vdash t : T$, then $t \rightarrow^*_{coq} v$, where v is a value of T*

Fact 2 *There is no value of type* `False`.

We have reached contradiction, so we have "proved" our statement.

5 Automatically Solving the Generated Goals

6 Implementation/Architecture/Results

We implemented the translation methods in `stainless.verification.CoqEncoder.scala` and `stainless.verification.CoqExpression.scala`. The generated methods are written into temporary files, and coqc is invoked on them. For each function, a separate .v file is created, and each file contains all the dependencies of the function. For simplicity, we assumed a function depends on all ADT's and the functions it invokes directly or transitively. In order to speed verification and to eliminate a domino effect caused by one failing verification step, all obligations of the dependencies are admitted. The logic handling .v files can be found in `stainless.verification.CoqIO.scala`

In order to integrate our solution into the Stainless framework we replaced `VerificationChecker` with `CoqVerificationChecker`. It creates a verification condition for each method (file) and puts its result to the report. Because functions are verified individually, their verification can be canceled, or timeout can be defined for them. In order to activate coq verification, a `--coq` flag has to be passed to stainless.

To benchmark the translation and verification, it has been run on the Lists library with 10 minutes timeout (per function case). The results are summarized in the following table:

7 Conclusions

We have implemented methods to transform Stainless programs with recursive functions and other nontrivial constructs into Coq programs. We also defined tactics to automatically generate proof for most of the cases. Our tactics are able to perform case analyis, and recursive function inlining without infinite loops. The automatic proof generation handles the majority of the functions in the Lists library. We also integrated our solution into the Stainless framework.

Future work includes extending the set of supported Stainless elements, for example with mutually recursive functions. The tactics can be also enhanced further to handle the remainder of the library. There is also room for improvement with the speed of tactics, as for example the external set-solver from stdpp takes a lot of time to fail, thus slows down every case where recursive inlining is needed. Checking termination of recursive functions is also a possible enhancement.