Formalization of Rice’s Theorem over a Functional Language Model

Thiago Mendonça Ferreira Ramos* and Ariane Alves Almeida* and Mauricio Ayala-Rincón*†

Departments of †Mathematics and *Computer Science
Universidade de Brasília
Brasília D.F., Brazil

Abstract
Classical proofs of Rice’s Theorem assume the existence of a universal (Turing) machine and build a reduction from the problem of deciding whether a machine halts or not to the problem of separability of semantic properties of machines. This work presents a formalization in PVS of Rice’s Theorem over a computational model given as a class of partial recursive functions. The main differences with classical proofs are that the given formalization is developed over a functional programming model and that the proof does not depend on the undecidability of the Halting Problem being made directly without using any translation to or from other computational models. As corollaries, straightforward formalizations of the undecidability of the Halting Problem, functional equivalence problem, existence of fixed points problem and self-replication problem are obtained.

Keywords: Functional Programming Models, Automating Termination, Computability Theory, Halting Problem, Rice’s Theorem, Theorem Proving, PVS

1 Introduction
In computer science, it is well-known that all Turing complete models of computation have some expressiveness limits, this means that not all kind of problems can be solved and that programs over each of these models with the same or different semantics cannot be distinguished. One of these limits is given by Rice’s Theorem. This result says that it is impossible to build a program that decides a semantic predicate over other programs, unless the predicate is either the set of all programs or the empty set.

The formalization is performed in the proof assistant PVS, using as model of computation a variant of a first-order functional language called PVS0. In this language there are constants, only one variable, unary and binary built-in operators, if-then-else instructions and recursive calls. PVS0 was designed with the main aim of formalizing the correction of mechanisms to automate verification of termination in PVS, which makes of high interest the exploration of the computability theory of PVS0 as a model of computation. Indeed, PVS0 is applied to formalize the equivalence among different criteria of termination, including size change principle [13] termination criteria such as Manolios and Vroom’s Calling Context Graphs [14] and Avelar et al Matrix Weighted Graphs [3], as well as Turing termination [17] (that is the criterion applied for the PVS specification language), and Dependency Pairs termination [1][2]. The libraries for PVS0, which include equivalence proofs of these termination criteria, are available as part of the NASA LaRC PVS library.

The aim of this paper is related with a second relevant objective that is the study of PVS0 as a model of computation, for which it is required the formalization of its computational theory. In addition to verification of termination criteria over PVS0, it is important to formalize computability properties of the model to certify that it is indeed a reasonable and expressive model of computation. Previous work presented the formalization of the undecidability of the Halting Problem for the PVS0 model [5]. In that work this result was formalized for a language model different from the one used in the current paper, which allows only programs that consist of a unique (recursive) function, called here the single-function PVS0 model. The language model used in this paper is more realistic, accepting, similarly to functional specifications, a list of functions such that each function can call and be called by all functions in the list using their indices. This model is called the multiple-function PVS0 model (or just PVS0 when no confusion arise). The undecidability of the Halting Problem

©2019 Published by Elsevier Science B. V.
was formalized for the multiple-function model too, but this result is now obtained just as a corollary of the formalization of Rice’s theorem. The PVS0 language model allows a unique input/output type, passed as a parameter of the PVS theory, and for dealing with Rice’s Theorem the type of naturals is used. Basic operators including successor, greater than and projections, provide a Turing Complete model. The formalization assumes that there is a manner of computing the own number of Gödel of each PVS0 program, i.e., the recursion theorem. Furthermore, it follows Cantor’s diagonal argument, similarly to standard proofs of the undecidability of the Halting Problem and uncountability of real numbers. An alternative formalization approach could be based on the construction of a universal program for the PVS0 model, but it would increase the complexity of the formalization. Using such construction, the proof would require reducing the Halting Problem to the problem of separability of semantic properties of PVS0 programs, which is not the case of the current formalization. Thus, the current formalization depends only on the above mentioned assumption and does not uses undecidability of the Halting Problem. In textbooks, assumptions such as the existence of universal machines and programs as well as Gödelization of machines and programs are intuitively given without providing complete constructions. In order to rule out the assumption and get a full formalization, the above mentioned basic operators are required to guarantee Turing completeness and thus the recursion theorem.

After giving the semantics of the PVS0 model and the specifications of computable and partial recursive classes in Sections 2 and 3, resp., Section 4 explains the formalization of Rice’s Theorem. Section 5 discusses crucial results formalized as simple corollaries of Rice’s Theorem including the undecidability of the Halting Problem. Then, before concluding and discussing current and future work in Section 7, Section 6 discusses related work.

The formalization of Rice’s Theorem is available at

https://github.com/thiagomendoncaferreiraramos/ricestheorem

It requires the installation of the NASA library https://github.com/nasa/pvslib, which includes the PVS0 library for termination criteria and their equivalence over the PVS0 single-function model.

2 Semantics of PVS0 Programs

Expressions of the PVS0 functional language have the grammar below.

\[
expr ::= \text{cnst}(T) \mid \text{vr} \mid \text{op1}(N, expr) \mid \text{op2}(N, expr, expr) \mid \text{rec}(N, expr) \mid \text{ite}(expr, expr, expr)
\]

Above, \(T\) is an uninterpreted type over which PVS0 expressions are interpreted. Constants of type \(T\) have grammar \(\text{cnst}(T)\) and the symbol \(\text{vr}\) is the unique symbol of variable. \(\text{op1}\) and \(\text{op2}\) denote respectively unary and binary operators indexed by naturals. The symbol \(\text{rec}\) is for recursive calls and is also indexed by naturals used to select the function in a PVS0 program to be called, applying it to the result of the evaluation of the second argument, \(expr\). Finally, \(\text{ite}\) is the symbol of the branching instruction and its evaluation has the same semantics as the instruction if-the-else.

The kernel of a PVS0 program is a non empty list of PVS0 expressions whose main expression (function) is the one in the head of the list and, for which the symbol of recursive calls, \(\text{rec}(i, \_\_\_\_\_\_\_\_\_\_\_)\), interprets functions calls as evaluations of the \(i\)th PVS0 expression in the list. PVS0 programs include also lists of unary and binary functions to interpret the symbols of unary and binary operators, and an element of the worked type to be interpreted as false (for the evaluation of the guards of \(\text{ite}\) instructions). Thus, PVS0 programs are 4-tuples of the form \((O_1, O_2, \perp, E_f)\), where \(O_1\) and \(O_2\) are the lists of unary and binary functions, \(\perp\) is an element of \(T\) to be interpreted as false, and \(E_f\) is the kernel of the program.

The (eager) evaluation predicate \(\varepsilon\), for PVS0 programs is defined below.
\[ \varepsilon(O_1, O_2, \bot, E_f)(e, v_i, v_o) := \text{CASES } \varepsilon \text{ OF} \]

\begin{align*}
\text{cnst}(v) & : v_o = v; \\
\text{vr} & : v_o = v; \\
\text{op1}(j, e_1) & : \exists (v' : T) : \\
& \quad \varepsilon(O_1, O_2, \bot, E_f)(e_1, v_i, v') \land \\
& \quad \text{IF } j < |O_1| \text{ THEN} \\
& \quad v_o = O_1(j)(v'); \\
& \quad \text{ELSE } v_o = \bot \\
\text{op2}(j, e_1, e_2) & : \exists (v', v'' : T) : \\
& \quad \varepsilon(O_1, O_2, \bot, E_f)(e_1, v_i, v') \land \\
& \quad \varepsilon(O_1, O_2, \bot, E_f)(e_2, v_i, v'') \land \\
& \quad \text{IF } j < |O_2| \text{ THEN} \\
& \quad v_o = O_2(j)(v', v''); \\
& \quad \text{ELSE } v_o = \bot \\
\text{rec}(j, e_1) & : \exists (v' : T) : \varepsilon(O_1, O_2, \bot, E_f)(e_1, v_i, v') \land \\
& \quad \text{IF } j < |E_f| \text{ THEN} \\
& \quad \varepsilon(O_1, O_2, \bot, E_f)(E_f(j), v', v_o) \\
& \quad \text{ELSE } v_o = \bot \\
\text{ite}(e_1, e_2, e_3) & : \exists (v' : T) : \varepsilon(O_1, O_2, \bot, E_f)(e_1, v_i, v') \land \\
& \quad \text{IF } v' \neq \bot \text{ THEN } \varepsilon(O_1, O_2, \bot, E_f)(e_2, v_i, v_o) \\
& \quad \text{ELSE } \varepsilon(O_1, O_2, \bot, E_f)(e_3, v_i, v_o). 
\end{align*}

Parameters \( v_i \) and \( v_o \) of \( \varepsilon \) are the input and output values. Note that the parameter \( e \) is needed since the evaluation of a program \((O_1, O_2, \bot, E_f)\), for short denoted as \( \text{pvs0} \), leads to evaluation of sub expressions \( \varepsilon \) of the expressions in \( E_f \) and, when a recursive call, as \( \text{rec}(j, e_1) \) in the recursive case above, is evaluated, the expression \( E_f(j) \) should be considered. Using the \( \varepsilon \) predicate, another predicate is defined that holds for \( \text{pvs0} \) programs and correct inputs and outputs, specified as below, where \( \text{pvs0}'4 \) denotes the fourth element of \( \text{pvs0} \).

\[ \gamma(\text{pvs0})(v_i, v_o) = \varepsilon(\text{pvs0})(\text{pvs0}'4(0), v_i, v_o) \]

Note that \( \gamma \) starts the evaluation from the main function of the program that is \( \text{pvs0}'4(0) \), the head of the list of expressions \( \text{pvs0}'4 \), the same as \( E_f \).

In order to show results related with automation of termination, semantic termination of \( \text{pvs0} \) programs was specified in [5]. This definition is required to prove completeness and equivalence of practical termination criteria as well as to formalize computability results such as undecidability of the Halting Problem and Rice’s theorem. Termination is specified as the semantic termination predicate below.

\[ T_e(\text{pvs0}, v_i) := \exists (v_o : T) : \varepsilon(\text{pvs0})(\text{pvs0}'4(0), v_i, v_o). \]

This predicate states that for a given program \( \text{pvs0} \) and input \( v_i \), the evaluation of the program’s expression \( \text{pvs0}'4(0) \) on the value \( v_i \) terminates with the output value \( v_o \). The program \( \text{pvs0} \) is total with respect to \( \varepsilon \) if it satisfies the following predicate (here notice that polymorphism in PVS allows the use of the same function or predicate name with different types).

\[ T_e(\text{pvs0}) := \forall (v : T) : T_e(\text{pvs0}, v). \]

Both the formalizations of the undecidability of the Halting Problem as given in [5] and Rice’s Theorem in this paper require to build composition of functions and programs to give rise to contradictions. Using the multiple-function \( \text{pvs0} \) model in this work, it is possible to specify the composition of arbitrary \( \text{pvs0} \) functions, but for the single-function model used in [5] it was not the case. The required compositions in [5] were specified manually as new single-function \( \text{pvs0} \) programs, which was possible since the specific functions to be composed were always terminating.
To compose PVS0 programs that share the same unary and binary operators and interpretation of false, a fundamental issue is the notion of offset used to adjust the indices of function calls in program lists. It works as a simplified version of offset in assembly languages, where instruction labels in some piece of code are adjusted when they are shifted. For PVS0 programs this is specified as the function $\beta$.

$$\beta(n)(e) := \text{CASES } e \text{ OF}$$

$$\text{cnst}(v) : \text{cnst}(v);$$

$$\text{vr} : \text{vr} ;$$

$$\text{op1}(j, e_1) : \text{op1}(j, \beta(n)(e_1));$$

$$\text{op2}(j, e_1, e_2) : \text{op2}(j, \beta(n)(e_1), \beta(n)(e_2));$$

$$\text{rec}(j, e_1) : \text{rec}(j + n, \beta(n)(e_1));$$

$$\text{ite}(e_1, e_2, e_3) : \text{ite}(\beta(n)(e_1), \beta(n)(e_2), \beta(n)(e_3))$$

Using $\beta$, the composition of $(O_1, O_2, \bot, A)$ and $(O_1, O_2, \bot, B)$, two PVS0 programs, is expressed by the following property.

$$\forall (v_1, v_0) : \exists (v) : \gamma(O_1, O_2, \bot, B)(v_1, v) \land \gamma(O_1, O_2, \bot, A)(v, v_0) \iff$$

$$\gamma(O_1, O_2, \bot, [\text{rec}(1, \text{rec}(1 + |A|, \text{vr})]) : \text{map}(\beta(1))(A) : \text{map}(\beta(1 + |A|))(B))(v_1, v_0)$$

As an example, consider the unitary lists of unary and binary operators below for predecessor and multiplication on $\mathbb{N}$, and element to interpret as false.

- $O_1 := \lambda(n : \mathbb{N}^+) : n - 1$
- $O_2 := \lambda(n, m : \mathbb{N}) : n \times m$
- $\bot := 0$

Consider now the following unitary lists of expressions, where the first one specifies the basic function quadratic and the second one the factorial function.

- quadratic := [op2(0, vr, vr)]
- factorial := [ite(vr, op2(0, vr, rec(0, op1(0, vr))), cnst(1))]

The 4-tuples $(O_1, O_2, \bot, \text{quadratic})$ and $(O_1, O_2, \bot, \text{factorial})$ are then PVS0 programs for these functions. The correction of the construction of the program for the composition of factorial and quadratic specified using $\beta$, is expressed as the property below.

$$\forall (v) : \gamma(O_1, O_2, \bot, [\text{rec}(1, \text{rec}(1 + |\text{factorial}|, \text{vr})]) :$$

$$\text{map}(\beta(1))(\text{factorial}) :$$

$$\text{map}(\beta(1 + |\text{factorial}|))(\text{quadratic}))(v_1, (v_2^2)!)$$

Since the result of factorial is 1, map($\beta(1))(\text{factorial})$ is ite(vr, op2(0, vr, rec(1, op1(0, vr))), cnst(1)) and map($\beta(1 + |\text{factorial}|))(\text{quadratic})$ is op2(0, vr, vr), the simplification of the 4-tuple written above is:

$$(O_1, O_2, \bot, [\text{rec}(1, \text{rec}(2, \text{vr}))] : [\text{ite}(\text{vr}, \text{op2}(0, \text{vr}, \text{rec}(1, \text{op1}(0, \text{vr})), \text{cnst}(1)))) : [\text{op2}(0, \text{vr}, \text{vr})]$$

A functional alternative for the semantic evaluation predicate $\varepsilon$ should take into consideration the case in which the evaluation does not return an output. The problem is solved by adding an element to the working type that is interpreted as none and denoted by $\Diamond$. In addition, the evaluation function includes a parameter that limits the allowed number of nested recursive calls: when this limit is reached, the function returns $\Diamond$. This is specified as the function $\chi$ below.
\( \chi(O_1, O_2, \bot, E_f) (n, e, v_i) := \) IF \( n = 0 \) THEN \( \diamond \) ELSE CASES \( e \) OF \\
\text{\texttt{cnst}}(v) : v;
\text{\texttt{vr}} : v_i;
\text{\texttt{op1}}(j, e_1) : IF \( j < |O_1| \) THEN \\
\text{LET} \( v' = \chi(O_1, O_2, \bot, E_f) (n, e_1, v_i) \) IN \\
IF \( v' = \diamond \) THEN \( \diamond \) ELSE \( O_1(j)(v') \)
ELSE \( \diamond \);
\text{\texttt{op2}}(j, e_1, e_2) : IF \( j < |O_2| \) THEN \\
\text{LET} \( v' = \chi(O_1, O_2, \bot, E_f) (n, e_1, v_i) \), \\
\( v'' = \chi(O_1, O_2, \bot, E_f) (n, e_2, v_i) \) IN \\
IF \( v' = \diamond \) OR \( v'' = \diamond \) THEN \( \diamond \) ELSE \( O_2(j)(v', v'') \)
ELSE \( \diamond \);
\text{\texttt{rec}}(j, e_1) : LET \( v' = \chi(O_1, O_2, \bot, E_f) (n, e_1, v_i) \) IN \\
IF \( v' = \diamond \) THEN \( \diamond \)
ELSIF \( j < |E_f| \) THEN \( \chi(O_1, O_2, \bot, E_f) (n - 1, E_f(j), v') \)
ELSE \( \bot \);
\text{\texttt{ite}}(e_1, e_2, e_3) : LET \( v' = \chi(O_1, O_2, \bot, E_f) (n, e_1, v_i) \) IN \\
IF \( v' = \diamond \) THEN \( \diamond \)
ELSIF \( v' \neq \bot \) THEN \( \chi(O_1, O_2, \bot, E_f) (n, e_2, v_i) \)
ELSE \( \chi(O_1, O_2, \bot, E_f) (n, e_3, v_i) \).

The predicate \( \varepsilon \) and function \( \chi \) are proved equivalent in the following sense:

\( \forall (v_1, v_2) : \varepsilon(peso)(v_1, v_2) \leftrightarrow \exists (n) : \chi(peso)(n, peso'4(0), v_1) = v_2 \land v_2 \neq \diamond \)

Similarly to [5], a terminating program, \( \text{\texttt{peso}} \), satisfies the predicate:

\( \forall (v_i) : \exists (v_o) : \gamma(peso)(v_i, v_o) \) or equivalently \( \forall (v_i) : \exists (n) : \chi(peso)(n, peso'4(0), v_i) \neq \diamond \)

To define the classes of partial recursive and computable functions, indices of the function calls in PVS0 programs are restricted to valid indices:

\( \text{valid\_index\_rec}(e, i) := \forall (i, e_1) : \text{subterm}(\text{\texttt{rec}}(i, e_1), e) \Rightarrow i < n \)

\( \text{valid\_index}(E_f) := \forall (i < |E_f|) : \text{valid\_index\_rec}(E_f(i), |E_f|) \)

### 3 PVS0 Computable and Partial Recursive Classes

Let \( O_1, O_2 \) and \( \bot \) be fixed unary, binary operators and false symbol and, besides that, fix the input and output type as naturals. Below, it is defined a class of partial recursive functions.

\( \text{\texttt{partial\_recursive}}(\text{\texttt{peso}}) := \text{\texttt{peso’1}} = O_1 \land \text{\texttt{peso’2}} = O_2 \land \text{\texttt{peso’3}} = \bot \land \text{valid\_index}(\text{\texttt{peso’4}}) \)

If \( \text{\texttt{partial\_recursive}}(\text{\texttt{peso}}) \) holds, \( \text{\texttt{peso}} \) is of type \textbf{PartialRecursive}. If in addition \( \text{\texttt{peso}} \) is terminating it is of type \textbf{Computable}.

To prove Rice’s Theorem, it is necessary to formalize some lemmas about code shifting. The first lemma is:

\( \forall (O_1, O_2, \bot, A, B, e, v_i, n) : \\
\chi(O_1, O_2, \bot, B)(e, v_i, n) = \chi(O_1, O_2, \bot, A :: \text{map}(\beta([A]))(B))(\beta([A])(e), v_i, n) \)
The second lemma is:

\[ \forall (O_1, O_2, s, e, A, n) : \forall (e \mid valid\_index\_rec(e, \{ A \})) : \forall (A \mid valid\_index(A)) : \chi(O_1, O_2, s, e, A, n) = \chi(O_1, O_2, s, e, \{ A \}) \]

Both lemmas are proved by induction on the lexicographical order given by pairs \((n, e)\) (of type \(\mathbb{N} \times \text{PVS0Expr}\)) built with the orders on naturals and (sub)expressions. They are used respectively to prove the next two lemmas, along with the equivalence between predicate \(\varepsilon\) and function \(\chi\).

\[
\forall (O_1, O_2, s, e, A, n) : \\
\varepsilon(O_1, O_2, s, e, A, n) \iff \varepsilon(O_1, O_2, s, e, A, n) \\
\forall (O_1, O_2, s, e, A, n) : \forall (e \mid valid\_index\_rec(e, \{ A \})) : \forall (A \mid valid\_index(A)) :
\]

Formalizing Rice's Theorem also requires a definition of \textit{semantic predicate} of programs and a \textit{Gödelization} of the partial recursive class of PVS0 programs.

The notion of semantic predicate over PVS0 programs is specified as:

\[
is\_semantic\_predicate(P) := \forall (\text{pvs0}_1, \text{pvs0}_2) : \\
(\forall (i, o) : \gamma(\text{pvs0}_1)(i, o) \iff \gamma(\text{pvs0}_2)(i, o)) \Rightarrow (P(\text{pvs0}_1) \iff P(\text{pvs0}_2))
\]

For the Gödelization, it is necessary to define a mapping from each PVS0 expression that works with naturals and a mapping from lists of naturals to naturals. The mapping from expressions to naturals uses a bijection from pairs of naturals to naturals, and both are defined as below.

\[
tuple2nat(m, n) := \frac{(m + n + 1)(m + n)}{2} + n
\]

\[
k_e(expr) := \text{CASES} expr \text{ OF}
\[
\text{vr} : 0; \\
\text{cns}(v) : v \times 5 + 1; \\
\text{rec}(j, e_1) : \text{tuple2nat}(j, \kappa_e(e_1)) \times 5 + 2; \\
\text{op1}(j, e_1) : \text{tuple2nat}(j, \kappa_e(e_1)) \times 5 + 3; \\
\text{op2}(j, e_1, e_2) : \text{tuple2nat}(j, \text{tuple2nat}(\kappa_e(e_1), \kappa_e(e_2))) \times 5 + 4; \\
\text{ite}(e_1, e_2, e_3) : \text{tuple2nat}(\kappa_e(e_1), \text{tuple2nat}(\kappa_e(e_2), \kappa_e(e_3))) \times 5 + 5;
\]

The bijective function from lists of naturals to naturals is called \(\alpha\).

Using the above notions, the injective function from the class of partial recursive functions to naturals is defined as below.

\[
\kappa_\text{p}(\text{pvs0}) = \alpha(\text{map}(\kappa_e)(\text{pvs0}^4))
\]

Finally, the next assumption is required.

\[
\forall(E_f) : \exists(\text{print} : \text{PartialRecursive}) : \\
\text{LET} \ self = (O_1, O_2, s, E_f :: \text{map}(\beta(|E_f|))(\text{print}^4)) \ 	ext{IN} \\
\text{partial}\_\text{recursive}(\text{self}) \land \forall i : \varepsilon(\text{self})(\beta(|E_f|))(\text{print}^4(0), i, \kappa_\text{p}(\text{self}))
\]

This assumption means that for each PVS0 program list there exists a partial recursive PVS0 program such that they both can be used to build another partial recursive program that outputs the own number of Gödel. As mentioned in the introduction, this result is known as the recursion theorem and, as for universal machines and programs, it is usually given in textbooks in an intuitive manner to convince the reader, but
without providing explicit constructions. In Turing complete models, it is possible to design entities that print
d Themselves. From this assumption, depending on the chosen lists of unary and binary functions, if there exists
the possibility of creating from a list of PVSo expressions a partial recursive PVSo program such that its output
for any evaluation is itself, then Rice’s theorem holds.

4 Formalization of Rice’s Theorem

Rice’s theorem, i.e. that any semantic predicate can be decided if and only if it is the set of all PVSo programs
or the empty set, is specified as below.

\[ \exists ( \text{decider} : \text{Computable}) : \]
\[ \forall (\text{pvso} : \text{PartialRecursive}) : \neg(\gamma(\text{decider})(\kappa_{p}(\text{pvso}), \bot) \iff P(\text{pvso})) \]
\[ \iff (P = \text{fullset} \lor P = \emptyset) \]

**Proof. Necessity**: Suppose that \( P = \text{fullset} \). Let \( T \) be an element different from \( \bot \). The PVSo program \( \text{decider} = (O_1, O_2, \bot, [\text{cnst}(T)]) \), decides \( \text{fullset} \). Now, suppose that \( P = \emptyset \). The PVSo program \( \text{decider} = (O_1, O_2, \bot, [\text{cnst}(\bot)]) \) decides \( \emptyset \).

**Sufficiency**: Proved by contraposition. Let assume that \( (P \neq \text{fullset} \land P \neq \emptyset) \). This implies that there are PVSo programs, say \( p \) and \( np \), such that \( P(p) \) and \( \neg P(np) \). For reaching a contradiction, suppose that there exists \( \text{decider} : \text{Computable} \) such that:

\[ \forall (\text{pvso} : \text{PartialRecursive}) : \neg(\gamma(\text{decider})(\kappa_{p}(\text{pvso}), \bot) \iff P(\text{pvso})) \]

And, consider the program \( \text{opp} \) with kernel:

\[ \text{opp} = \left[ \text{ite}(\text{rec}(1, \text{rec}(1 + |\text{decider}'| + |np'| + |p'|, \text{vr})), \right. \]
\[ \left. \text{rec}(1 + |\text{decider}'|, \text{vr}), \text{rec}(1 + |\text{decider}'| + |np'|, \text{vr})] : \right. \]
\[ \text{map}(\beta(1))(\text{decider}') :: \]
\[ \text{map}(\beta(1 + |\text{decider}'|))(np') :: \]
\[ \text{map}(\beta(1 + |\text{decider}'| + |np'|))(p') \]

Using the assumption that there are programs in the model that can print their own number of Gödel, making
\( E_f = \text{opp} \):

\[ \exists (\text{print} : \text{PartialRecursive}) : \]
\[ \text{LET } \text{self} = (O_1, O_2, \bot, \text{opp} :: \text{map}(\beta(\text{opp}))(\text{print}')) \text{ IN} \]
\[ \text{partial.recursive}(\text{self}) \land \forall i : \varepsilon(\text{self})(\beta(|\text{opp}||(\text{print}'|0)), i, \kappa_{p}(\text{self})) \]

To understand how \( \text{opp} \) and \( \text{self} \) work, suppose that for each PVSo program \( \text{PartialRecursive} \) there is a
function with the same name; in pseudo-code:

\[ \text{self}(n) := \text{IF } \text{decider}(\kappa_{p}(\text{self})) \neq \bot \text{ THEN } np(n) \text{ ELSE } p(n) \]

The proof uses Cantor’s diagonal argument. If \( \text{decider}(\kappa_{p}(\text{self})) \neq \bot \), then \( P(\text{self}) \), but \( \text{self} \) behaves as \( np \) and thus \( \neg P(\text{self}) \) holds, which is a contradiction. Otherwise, if \( \text{decider}(\kappa_{p}(\text{self})) = \bot \), then \( \neg P(\text{self}) \), but \( \text{self} \) behaves as \( p \) and thus \( P(\text{self}) \) that is a contradiction too. This is the main idea behind the rest of the
explanation of the formalization.

The aforementioned assumption implies that there exists an element of the partial recursive class, say \( \text{print} \),
such that:

\[ \text{LET } \text{self} = (O_1, O_2, \bot, \text{opp} :: \text{map}(\beta(\text{opp}))(\text{print}')) \text{ IN} \]
\[ \text{partial.recursive}(\text{self}) \land \forall i : \varepsilon(\text{self})(\beta(|\text{opp}||(\text{print}'|0)), i, \kappa_{p}(\text{self})) \]
Making $p_vso = self$ it can be concluded that

$$\neg\gamma(decider)(\kappa_p(self), \bot) \Leftrightarrow P(self)$$

The proof splits into two sub-cases.

**Sub-case 1:** $P(self)$. In this case, $\neg\gamma(decider)(\kappa_p(self), \bot)$ is concluded.

Since $P$ is a semantic predicate, one has:

$$\forall(p_vso_1, p_vso_2) :$$

$$(\forall(i, o) : \gamma(p_vso_1)(i, o) \Leftrightarrow \gamma(p_vso_2)(i, o)) \Rightarrow (P(p_vso_1) \Leftrightarrow P(p_vso_2))$$

Thus, choosing $p_vso_1$ as $self$ and $p_vso_2$ as $np$, it gives:

$$(\forall(i, o) : \gamma(self)(i, o) \Leftrightarrow \gamma(np)(i, o)) \Rightarrow (P(self) \Leftrightarrow P(np))$$

Assuming $\forall(i, o) : \gamma(self)(i, o) \Leftrightarrow \gamma(np)(i, o)$, by $P(self)$, $P(np)$ also holds, which is a contradiction since $\neg P(np)$.

Consequently, $\neg\forall(i, o) : \gamma(self)(i, o) \Leftrightarrow \gamma(np)(i, o)$ should hold.

But this is not possible because $self$ performs the same as $np$ as showed below.

Starting by $\gamma(self)(i, o)$ and expanding $\gamma$, and from $\varepsilon(self)(self'4(0), i, o)$ replacing $self$ by its definition, one obtains:

$$\varepsilon(self)(opp :: map(\beta(opp))(0), i, o)$$

That by properties of lists and definition of $opp$ gives $\varepsilon(self)(opp(0), i, o)$, and then:

$$\varepsilon(self)(ite(rec1, rec1(1 + |decider'4| + |np'4| + |p'4|, vr)), rec1(1 + |decider'4|, vr), rec1(1 + |decider'4| + |np'4|, vr), i, o)$$

Then, by the definition of $\varepsilon$ and operational semantics of $ite$, one has:

$$\exists(v') : \varepsilon(self)(rec1, rec1(1 + |decider'4| + |np'4| + |p'4|, vr), i, v') \land$$

IF $v' \neq \bot$ THEN $\varepsilon(self)(rec1 + |decider'4|, vr), i, o)$

ELSE $\varepsilon(self)(rec1 + |decider'4| + |np'4|, vr, i, o)$

Further, by adequate expansions of predicate $\varepsilon$ and application of equalities $self'4(1) = \beta(1)(decider'4(0))$, and $self'4(1 + |decider'4| + |np'4| + |p'4|) = \beta(|opp|)(print'4(0))$, one has:

$$\exists(v') : \exists(v'') : \exists(v''' : i = v''' \land$$

$$\varepsilon(self)(\beta(|opp|)(print'4(0)), v''', v''') \land$$

$$\varepsilon(self)(\beta(1)(decider'4(0)), v'', v') \land$$

IF $v' \neq \bot$ THEN $\varepsilon(self)(rec1 + |decider'4|, vr), i, o)$

ELSE $\varepsilon(self)(rec1 + |decider'4| + |np'4|, vr, i, o)$

And then, by Skolemization of the existentially quantified variables one has:

$$i = v''' \land$$

$$\varepsilon(self)(\beta(|opp|)(print'4(0)), v''', v''') \land$$

$$\varepsilon(self)(\beta(1)(decider'4(0)), v'', v') \land$$

IF $v' \neq \bot$ THEN $\varepsilon(self)(rec1 + |decider'4|, vr), i, o)$

ELSE $\varepsilon(self)(rec1 + |decider'4| + |np'4|, vr, i, o)$

And then, by Skolemization of the existentially quantified variables one has:
By the assumption, $\forall i : \epsilon(self)(\beta(\text{opp})(\text{print}'4(0)), i, \kappa_p(self))$, and instantiating $i = v''$ one obtains:

\[
\epsilon(self)(\beta(\text{opp})(\text{print}'4(0)), v'', \kappa_p(self)) \land \\
\epsilon(self)(\beta(\text{opp})(\text{print}'4(0)), v'', v'') \land \\
\epsilon(self)(\beta(1)(\text{decider}'4(0)), v'', v') \land \\
\text{IF } v' \neq \bot \text{ THEN } \epsilon(self)(\text{rec}(1 + |\text{decider}'4|, \text{vr}), i, o) \\
\text{ELSE } \epsilon(self)(\text{rec}(1 + |\text{decider}'4| + |np'4|, \text{vr}), i, o)
\]

Since the relation $\epsilon$ (is formalized to be) functional, one has that $v'' = \kappa_p(self)$. Thus, the expression below should be considered.

\[
\epsilon(self)(\beta(1)(\text{decider}'4(0)), \kappa_p(self), v') \Leftrightarrow \epsilon(\text{decider})(\text{decider}'4(0), \kappa_p(self), v')
\]

Thus one obtains,

\[
\epsilon(\text{decider})(\text{decider}'4(0), \kappa_p(self), v') \land \\
\text{IF } v' \neq \bot \text{ THEN } \epsilon(self)(\text{rec}(1 + |\text{decider}'4|, \text{vr}), i, o_1) \\
\text{ELSE } \epsilon(self)(\text{rec}(1 + |\text{decider}'4| + |np'4|, \text{vr}), i, o)
\]

By the hypothesis of this case, one has $\neg\gamma(\text{decider})(\kappa_p(self), \bot)$ that means that $v' \neq \bot$. Thus,

\[
\epsilon(self)(\text{rec}(1 + |\text{decider}'4|, \text{vr}), i, o)
\]

By adequate expansions of predicate $\epsilon$, Skolemization of the obtained existentially quantified variable as $v'_1$ and replacing the necessary variables, one obtains:

\[
\epsilon(self)(\text{np}'4(0), i, o)
\]

Applying lemmas of code shift:

\[
\epsilon(\text{np})(\text{np}'4(0), i, o)
\]

That is equivalent to $\gamma(\text{np})(i, o)$. Thus $\neg\forall(i, o) : \gamma(self)(i, o) \Leftrightarrow \gamma(\text{np})(i, o)$ does not hold and that is a contradiction.

**Sub-case 2: $\neg P(self)$.** It follows analogously to sub-case 1, except for the supposition that $P$ is a semantic predicate where $\text{pvsO}_1$ and $\text{pvsO}_2$ are instantiated respectively as $self$ and $p$, which gives a contradiction.

\[\square\]

## 5 Applications

Generality of Rice’s Theorem allows simple formalizations of important undecidability results in computability theory. In particular, since our proof does not depend on the undecidability of the Halting Problem, we obtain it as a direct consequence.

**Corollary 5.1 (Undecidability of the Halting Problem)**

\[
\neg\exists(\text{decider} : \text{Computable}) : \\
\forall(\text{pvsO} : \text{PartialRecursive}) : \neg(\gamma(\text{decider})(\kappa_p(\text{pvsO}), \bot) \Leftrightarrow T_\epsilon(\text{pvsO}))
\]

**Proof.** The formalization uses Rice’s Theorem instantiating the semantic predicate as $T_\epsilon$. Since the set $T_\epsilon$ is neither equal to the empty set nor to the whole set $\text{PartialRecursive}$, there exists no computable decider
Corollary 5.2 (Undecidability of Existence of Fixed Points)

\[ \neg \exists (\text{decider} : \text{Computable}) : \forall (\text{pesso} : \text{PartialRecursive}) : \neg (\gamma (\text{decider})(\kappa_\text{p}(\text{pesso}), \bot) \iff \exists (p) : \gamma (\text{pesso})(p, p)) \]

Proof.

The formalization instantiates Rice's Theorem using the semantic predicate

\[ \lambda (\text{pesso} : \text{PartialRecursive}) : \exists (p) : \gamma (\text{pesso})(p, p) \]

The predicate is then shown to be different from the empty set and from the whole set PartialRecursive. Indeed, on one side, the predicate holds for the program \((O_1, O_2, \bot, [\text{cnst}(0)]\)), showing that it is different from empty set. On the other side, it does not hold for the program \((O_1, O_2, \bot, [\text{rec}(0, \text{vr})])\) does not.

Corollary 5.3 (Undecidability of Self Replication)

\[ \neg \exists (\text{decider} : \text{Computable}) : \forall (\text{pesso} : \text{PartialRecursive}) : \neg (\gamma (\text{decider})(\kappa_\text{p}(\text{pesso}), \bot) \iff \exists (p : \text{PartialRecursive}) : \forall (i) : \gamma (p)(i, \kappa_\text{p}(p)) \land \gamma (\text{pesso})(i, \kappa_\text{p}(p))) \]

Proof.

To formalize it, it is necessary to instantiate the predicate in the Rice's theorem as

\[ \lambda (\text{pesso} : \text{PartialRecursive}) : \exists (p : \text{PartialRecursive}) : \forall (i) : \gamma (p)(i, \kappa_\text{p}(p)) \land \gamma (\text{pesso})(i, \kappa_\text{p}(p)) \]

The next step is showing that the predicate is neither the empty set nor the full PartialRecursive set. Using the assumption of the recursion theorem and instantiating it with \([\text{rec}(1, \text{vr})]\), one shows that the predicate is not empty. On the other side, the program \((O_1, O_2, \bot, [\text{op2}(i, \text{vr}, \text{cnst}(1))]\)) shows that the predicate is not the whole PartialRecursive set.

Corollary 5.4 (Undecidability of Functional Equivalence)

\[ \neg \exists (\text{decider} : \text{Computable}) : \forall (\text{pesso}_0, \text{pesso}_1 : \text{PartialRecursive}) : \neg (\gamma (\text{decider})(\text{tuple2nat}(\kappa_\text{p}(\text{pesso}_0), \kappa_\text{p}(\text{pesso}_1)), \bot) \iff \forall (i, o) : \gamma (\text{pesso}_0)(i, o) \iff \gamma (\text{pesso}_1)(i, o)) \]

Proof.

Suppose that there exists a Computable program decider that decides the above equivalence between functions. Then, instantiate \text{pesso}_0 above as the constant zero program \((O_1, O_2, \bot, [\text{cnst}(0)]\)). Then, the problem functional equivalence problem is reduced to decide if a program preforms the same as the constant zero program. The next step is instantiating the Rice’s Theorem with the following semantic predicate:
To show that the predicate is neither the empty nor the full $\text{PartialRecursive}$ set, it is enough to prove that the constant zero program belongs to the predicate and that the constant one program does not. After that, one uses the assumed program $\text{decider}$ to build another program for the equivalence with the constant zero program; this program is build as

$\lambda(pvso : \text{PartialRecursive}) : \forall(i, o) : \gamma(O_1, O_2, \bot, [\text{cnst}(0)])(i, o) \Leftrightarrow \gamma(pvso)(i, o)$

where $i$ is the index to the $\text{tuple2nat}$ function. Using the shifting code lemmas, it can indeed be simplified to decide equivalence to the constant zero program.

This formalization, requires also proving that the program built above is in fact $\text{Computable}$. This is a consequence of $\text{decider}$ being assumed as a $\text{Computable}$ program and then being terminating too. The proof concludes by applying again the shifting code lemmas and to show that the program built above is also terminating.

6 Related Work

Nowadays, mechanical proofs of computability properties is not only of interest as an exercise of formalization, but also of great importance to provide formal support for computational models used for pragmatical issues. As mentioned in the introduction, the main aim of the PVS0 computational single- and multiple-function models is related with the development of automation mechanisms to verify termination of PVS programs. In [5], PVS0 programs not only consist of a single function, but also were constrained in inductive levels such that in the level zero only the basic functions successor, greater than and projections are allowed and, in subsequent levels, other $\text{Computable}$ functions can be specified allowing calls to functions of the previous level as operators. Building composition of such PVS0 programs was not straightforward, which makes difficult the formalization of results such as Turing completeness and Rice’s Theorem. As seen in Section 2 on semantics of PVS0, the composition of programs specified in the multiple-function PVS0 language is straightforwardly achieved by application of the operator $\beta$.

For the single-function PVS0 language, the composition of two (non necessarily terminating) programs requires the construction of a third program which cannot be specified in a general manner since this depends on the (combinatorial) structure of the input programs. In the proof of undecidability of the Halting Problem in [5], this problem was easily resolved since only very particular composition constructions (of assumed terminating functions) were necessary. Such difficulties are solved in the current work using as model a language which supports specification of programs that consist of several functions that can call not only themselves recursively, but that can also call each other.

The equivalence between termination criteria was formalized for the single-function PVS0 model considered in [5]. However, it was not formalized completely for the current multiple-function PVS0 model. Theoretically, all termination criteria mentioned in the introduction (references [13], [14], [3], [17], [2]) work for both models, but technically, some of the criteria require a re-adaptation to deal simultaneously with static analysis of multiple-function programs that allow even mutual recursion (which, in particular, is avoided in the PVS functional specification language).

Computability properties have been formalized since the development of the first theorem provers and proof assistants. As well-known examples one can mention the mechanical proof of the undecidability of the Halting Problem in [4] using as model of computation the LISP language, as well as the formalization in Agda of the same theorem in [10] using as a model of computation axioms over the elements of an abstract type Prog.

Here, the focus is restricted to recent works in which computability results have been formalized over such computational models related with lambda calculus and programming languages. In [9], Foster and Smolka used as model of computation call-by-value lambda calculus, which is a Turing Complete model of computation, where beta-reduction can be applied whenever the beta-redex is not below an abstraction and the argument is an abstraction. For this model, the authors formalized several computational properties including Rice’s Theorem; indeed, they formalized that semantic predicates such that there are elements both in them and in their complements, are not recognizable. Using this result they concluded Rice’s Theorem. Also, Norrish formalized in [15], using HOL4, Rice’s Theorem for the model of lambda calculus, among others properties such as the existence of universal machines and an instance of the s-m-n theorem.

In addition to Rice’s Theorem, there are another computability results formalized over linguistic computational models. Forster, Kirsk and Smolka formalized in [7] undecidability of validity, satisfiability, and provability of first-order formulas following a synthetic approach based on the computation native to Coq’s constructive type theory. Forster and Larchey-Wendling formalized in [8] using Coq, the reduction of the Post Correspondence Problem via binary stack machines and Minsky machines to provability of intuitionistic linear logic. They started with Post Correspondence Problem and built a chain of reductions passing through binary
Post Correspondence Problem, binary Post Correspondence Problem with indices, binary stack machines, Minsky machines and finally provability of intuitionistic linear logic. Also, Larchey-Wendling formalized in [11] that any function of the type \( N^k \rightarrow N \) specified using Coq is total. Recently, previous author together with Forest formalized in [12] the undecidability of Hilbert’s Tenth Problem using a chain of reductions of problems: Halting Problem for TMs, Post’s Correspondence Problem, a specialized Halting Problem for Minsky Machines, FRACTAN (a language model that deals with register machines) termination, and solvability of Diophantine logic and of Diophantine equations.

7 Conclusions, Work in Progress and Future Work

A formalization of Rice’s Theorem in PVS is given over the functional language PVS0 seen as a model of computation. The formalization uses a Gödelization of PVS0 programs and follows Cantor’s diagonal argument to build a contradiction of the existence of a PVS0 program that can decide semantic predicates over PVS0 programs. Applications of this formalization include formalizations of corollaries for undecidability of the Halting Problem, functional equivalence problem, existence of fixed points problem and self-replication.

Other results of interest to be formalized for the PVS0 language model are the s-m-n theorem, the undecidability of Post Correspondence Problem, Post’s Theorem, Turing-completeness, the existence of a universal machine, the existence of self-replicate machines, linear speedup theorem, tape compression theorem, time hierarchy theorem, space hierarchy theorem, etc. The main difficulty, but also the interesting aspect of such formal developments in PVS0, is that the classical proofs of these theorems are performed over specific models such as lambda-calculus and Turing-machines. Even more interesting will be the formalization of other undecidability results outside the context of properties of computational models such as the Word Problem for algebraic structures ([16]) and Hilbert’s Tenth Problem ([12]).

Recent examples of related developments include the formalizations in Coq of the Post’s theorem for weak call-by-value lambda-calculus [9] and of the undecidability of the Post Correspondence Problem via reduction of the Halting Problem for Turing machines [6]. Despite the existence of correspondences between the functional model PVS0, lambda-calculus and Turing machines, which may be explored for the formalization of such theorems for PVS0 programs, obvious difficulties are that these formalizations are strongly related to the respective computational model and that (formally building) the required translations is not straightforward.

Since in the PVS0 library, equivalence results for termination criteria were formalized for the single-function model, it is also of great interest formalizing these results for the multiple-function model. There are different manners to achieve this objective such as: ensuring Turing Completeness for both models, and thus equivalence of the models; formalizing directly the equivalence between the models and building single-function to and from multiple-function program translations; or by directly adapting the results for the multiple-function model.

Related current work includes the formalization of the fixed-point theorem for the multiple-function PVS0 model. The proof of this theorem involved the assumption of the existence of a universal program (as mentioned before, a standard assumption in textbooks, but required to be constructed in complete formalizations), which was required to build a diagonal program that performs analogously to the lambda term \( \lambda xy \cdot (xx)y \).

References


