# On Light-Weight Verification and Heavy-Weight Testing\*

Friedrich W. von Henke<sup>1</sup>, Sam Owre<sup>2</sup>, Stephan Pfab<sup>1</sup>, Harald Rueß<sup>2</sup>

¹ Universität Ulm² SFakultät für InformatikCJames-Franck-Ring33D-89069 UlmM{vhenke,pfab}@informatik.uni-ulm.de{c

<sup>2</sup>SRI International Computer Science Laboratory 333 Ravenswood Ave. Menlo Park, CA, 94025 {owre,ruess}@csl.sri.com

Abstract. We give an overview on our approach to symbolic simulation in the PVS theorem prover and to demonstrate its usage in the realm of validation by executing specification on incomplete data. In this way, one can test executable models for a possibly infinite class of test vectors with one run only. One of the main benefits of symbolic simulation in a theorem proving environment is enhanced productivity by early detection of errors and increased confidence in the specification itself.

#### 1 Introduction

Traditional, simulation-based validation methods have not kept up with the scale or pace of modern digital design, and, therefore, form a major impediment in the drive for evermore complex designs [7]. This is mainly due to the sheer number of possible test cases which makes it nearly impossible to perform exhaustive testing. Thus testing only demonstrates the existence but not the absence of flaws. Even worse, it is unlikely that testing alone would have caught errors like the famous bug in the lookup table of the Intel Pentium floating-point division unit, since it only occurred on table inputs that were thought to be beyond the region of interest [15].

Formal verification methods based on theorem proving techniques, model-checking, or a combination thereof offer viable alternatives to simple testing, since formal methods permit proving the absence of errors (in the formal model). Unfortunately, the construction of formal justifications is usually at best semi-automatic for industrial-sized designs and the cost of doing formal analysis in an interactive way currently prevents formal methods from being integrated in the development cycle for both hardware and software.

In many cases, however, formal models are *executable* and amenable to validation by means of simulation. Such an execution facility is required to support evaluation of partially specified models including uninterpreted constants

<sup>\*</sup> This work has been partially supported by the Deutsche Forschungsgemeinschaft (DFG) project *Verifix* and by the Deutscher Akademischer Austauschdienst (DAAD). The work undertaken at SRI was partially supported by the National Science Foundation under grant No. CCR-9509931.

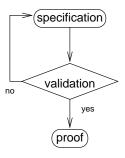


Fig. 1. Early Phase of Specification Life-Cycle

or function symbols, i.e., symbolic simulation, in order to be widely applicable. It provides a useful capability that is intermediate between running individual test cases and exploring properties for all runs, since symbolic simulation on indeterminate data permits covering a class of cases—possibly of infinite cardinality—with one test run. Put in other words, symbolic simulation may be regarded as a hybrid of light-weight verification and heavy-weight testing.

One of the main benefits of symbolic simulation in a theorem proving environment is enhanced productivity by early detection of errors and increased confidence in the specification, since symbolic evaluation permits investigating formal models fully automatically before resorting to formal proofs. This idealized view on the early phase of a specification life-cycle is depicted in Figure 1.

The applicability of symbolic simulation in the ACL2 [9] theorem prover has already been demonstrated in [10]. This system inherits efficient simulation in a natural way from the underlying implementation language, since its logic of recursive function is embedded in applicative Common Lisp. The situation is different for proof systems like Hol [5], Isabelle [13], Coq [4], or PVS [11] that are all based on powerful static type systems that are beyond the capabilities of type systems of current programming languages. The PVS [11] specification language, for example, relies on a rich set of typing constructs—like dependent types, predicate subtypes for expressing partial functions, and a powerful module system—in order to express formal models in a precise and succinct way. As a consequence, evaluation in systems with powerful static type systems cannot directly be inherited as in ACL2, and the reduction relation is usually implemented via substitution calculi. Although conceptually simple, these calculi are several orders of magnitude slower than specialized interpreters and compilers of off-the-shelf programming languages. This slowness has proved to be a formidable bottleneck for many medium- to large-size verification efforts such as the verification of pipelined microprocessors [18, 19].

The purpose of this paper is to give an overview on our approach of integrating efficient symbolic simulation in the PVS theorem prover. The basic approach is to provide translations between PVS functions and Common Lisp in order to

exploit the efficiency of Common Lisp for executing PVS functions. We demonstrate the usage of symbolic simulation in the realm of validating state-based models such as microprocessor specifications.

This paper is structured as follows. Section 2 describes our approach of symbolically executing PVS specifications. In order to make this paper largely self-contained, we recapitulate in Section 3 a basic method for encoding state machines in PVS; the running example is a model of a small, stack-based microprocessor model. Section 4 demonstrates the use of symbolic evaluation for early detection of errors. Finally, Section 5 closes with some remarks.

## 2 Efficient Symbolic Simulation

A certain subset of the PVS specification language [12] can be regarded as an executable, functional programming language that includes all kinds of operations on expressions of basic types, conditionals (if, cases, cond, table), total recursive functions defined by means of measure recursion, and structural recursion on abstract datatypes (catamorphisms, paramorphisms). Here we describe an extension to PVS for (symbolically) evaluating programs with Lisp-like speed.

We use the idea of inverse evaluation by Berger and Schwichtenberg [1] and compute a normal form for a functional PVS expression in three successive steps: first, an expression is translated into the corresponding (Common) Lisp program; second, the Lisp program is executed using Lisp's evaluation function eval; finally, the result of Lisp evaluation is translated back to a corresponding PVS expression. This includes the translation of Lisp closure to PVS λ-expressions by generating a bound PVS variable, by evaluating the closure on the Lisp translation of this variable, and by retranslating the result to the PVS level. In addition to the technique described in [1] we support evaluation of abstract datatype expressions and include the compilation of recursively defined PVS functions—which are required by the type system of PVS to be total. Altogether, we obtain Lisp-like execution speed for normalizing functional PVS expressions including uninterpreted constants and function symbols, since we can readily use Lisp compilers to produce efficient machine code for executing PVS functions.

Obviously, evaluation of programs that include uninterpreted constant and function symbols yield boolean conditions that evaluate to neither true nor false (e.g. IF (x + f(2) < 2 \* x) THEN  $e_1$  ELSE  $e_2$  ENDIF). In these cases we make use of the PVS prover to simplify or, whenever possible, to decide such formulas (including quantification). For the expressiveness of the full PVS logic, however, there cannot be a single proof procedure for deciding all kinds of formulas. Therefore, our animator is parameterized with respect to a prover strategy in order to simplify expressions in a problem-specific way. This functionality is realized in the Lisp compilation if\* of PVS conditionals.

Consider the simple example of evaluating fac(n + 2), where the factorial function fac is specified as a recursive PVS function and n is an unknown natural number.

```
(norm "fac(n + 2)" :strategy (assert))
```

The expression to be evaluated is presented as a string to the animator norm and the strategy argument (assert) causes this evaluator to use the PVS decision procedures to simplify conditionals. In the example above, the prover is—not too surprisingly—able to decide that  $n + 2 \le 1$  does not hold (since n is of type nat), but, without further information about n from the current context, one can not decide whether  $n + 1 \le 1$  is true; thus, evaluation stops at this point.

#### 3 State Machines

This section describes an often-used method for describing state machines in PVS (see also [17]). As a simple example we use a transcription of the stack machine from [3]. Furthermore, we depict the process of generating Lisp code for symbolically evaluating this machine.

The specification of the stack machine is packaged in a theory that is parameterized with respect to the length N of the memory array. All naturals less than N are valid addresses, the enumeration type opcodes in 1 lists the opcodes of the machine, and the record type instr determines the format of instructions.

```
address: TYPE = below[N]

opcodes: TYPE = {MOVE, MOVEI, MOVEWIND, MOVERIND, ADD, SUB, INCR, DECR, JUMP, JUMPZ, CALL, RET, HALT}

instr: TYPE = [# op: opcodes, arg1, arg2: address #]
```

States consist of a program counter, a stack, the memory, a flag for halting the processor, and the program code. Both the memory array and the program array are modeled as finite functions, and states are represented as elements of the record type state.

```
state: TYPE = [# pc: address,
stk: list[address],
mem: [address -> nat],
halted: bool,
code: [address -> instr]
#]
```

The five state components are accessed by pc, stk, mem, halted, and code, respectively.

Individual instructions are given semantics by defining state transformers that appropriately modify the current state s of the machine. Functions and records my be "modified" in PVS by means of an override expression. The result of an override expression is a function or record that is exactly the same as the original, except that at the specified arguments it takes the new values. For example, the move instruction with indirect addressing movewind applied to addresses a1, a2 increments the program counter (modulo N) and updates the memory at address mem(s)(a1) with the value mem(s)(a2) if the location to be updated is valid; otherwise the machine is halted.

```
| movewind(a1, a2: address)(s: state): state =
| IF mem(s)(a1) < N THEN
| s WITH[pc := inc(pc(s)),
| mem := mem(s) WITH[mem(s)(a1) := mem(s)(a2)]]
| ELSE
| s WITH [halted := TRUE]
| ENDIF
```

Recall that mem(s) is a function with codomain address. Likewise, movewind(a1, a2) is a (curried) function from states to states. Given such a function for every instruction, a one-step interpreter execute for the stackmachine is defined by case analysis on the opcode of the given instruction.

```
exec1(i: instr): [state -> state] =

LET a1 = arg1(i), a2 = arg2(i) IN

CASES op(i) OF

MOVE : move(a1, a2),

MOVEI : movei(a1, a2),

MOVEWIND: movewind(a1, a2),

...,

JUMP : goto(a1),

HALT : halt

ENDCASES
```

Using the techniques described in Section 2, the memory component mem(s), for example, is compiled to the following Lisp function. Note that here the memory size N is instantiated with 30.

If the argument of this function is a Lisp vector then one simply uses the Lisp vector lookup svref. Otherwise, the argument is uninterpreted and we compute a closure by wrapping a PVS application. The function phi retranslates Lisp terms (here: a Lisp symbol) to PVS expressions (here: a variable of type pvs::address[30]).

Similarly, the movewind function in 3 translates to Lisp code that is structurally similar to the corresponding PVS function.

Records are currently translated to Lisp vectors, the PVS  $\lambda$ -expression mem(s) is realized as a Lisp closure, memory lookup mem(s)(a) translates to the Lisp function application (funcall (mem s) a), and memory lookup on the Lisp level is simply encoded as function overwrite.

This naive translation scheme has the disadvantage that new state vectors are allocated in every simulation step of the machine and, even worse, the number of conditionals to be decided for array lookup depends on the number of updates of the array. Modern compilation technology such as structural analysis or monads [8] could be used to guarantee single-threadedness and, consequently, to use in-place updates in a safe way. Structural analysis, however, is non-trivial to implement and monads require a specialized specification style. Currently, we only support simple runtime tests to ensure safe in-place updates. More specifically, finite functions of type [below[n] -> A], where n is a positive integer and A an arbitrary type, are translated to a structure containing a tag and a vector for representing the finite function (array) under consideration. A second tag stored in the vector is used by the functions lookup and a destructive version of update to ensure safe in-place updates. The function update creates a new structure with a new tag and a shallow copy of the vector which is modified to hold the new data and the new tag. If the program is not single-threaded then one of lookup or update detects a mismatch of the tags and aborts at run-time. Consequently, using this approach it is safe to use in-place updates at the expense of some runtime overhead. The specification of the stack machine above, for example, is single-threaded, and the use of destructive updates yields run times that are several orders of magnitude faster than with the naive approach described above.

# 4 Symbolically Simulating State Machines

In this section some characteristic features of symbolic simulation in PVS are demonstrated using the running example of computing the minimum element in an array with the stack machine in the previous section. The semantics of this machine has been given in terms of a one-step interpreter, and the machine's basic cyclic behavior <code>exec(s, n)</code> is then defined by iterating the one-step interpreter <code>exec1</code> on the state s. The integer argument n serves hereby as an upper bound on the number of steps in order to guarantee termination (all functions in PVS are required to be (provably) total). Such an interpreter is formalized in the module <code>exec</code> in 5 in a machine-independent way. This module is parameterized with respect to a state type, a one-step interpreter, and a predicate for characterizing abort states. In addition, the parameter <code>observe</code> can be used to observe the dynamic behavior of the interesting parts of states.

```
5
exec[
    state: TYPE,
    step: [state -> state],
 halted?: pred[state],
        A: TYPE,
 observe: [state -> A]
]: THEORY
BEGIN
exec_rec(n: nat, s: state, acc: list[A]): RECURSIVE list[A] =
   IF n = 0 OR halted?(s) THEN reverse(acc)
    ELSE LET s1 = step(s), newacc = cons(observe(s1), acc) IN
            exec_rec(n - 1, s1, newacc) ENDIF
MEASURE n
exec(max: nat, s: state): list[A] =
   exec_rec(max, s, null)
END exec
```

The recursive function exec\_rec iterates the step function and accumulates the observable part of resulting states; the type system of PVS together with the MEASURE ensures that this function is total. One advantage of our approach of evaluating PVS functions on the Lisp level lies in the fact that a Common Lisp compiler can readily be used to eliminate, for efficiency reasons, tail-recursive calls like in exec\_rec.

A particular interpreter for the stack machine is obtained by instantiating the module with actual parameters. The fourth and the fifth parameter below indicate that the complete state is being observed.

```
sm: THEORY = exec[state, step, halted?,
state, (LAMBDA (s: state): s)]
```

The program MIN in 7 for computing the index of the minimum element in an array is used as a running example. Since the format of all instructions is fixed to contain exactly two arguments, an uninterpreted family X(i) of "don't cares" models unused argument positions. The loader function simply computes an array from the more convenient list notation of programs; its definition is not shown here.

```
7
X: [nat -> address]
MIN: ARRAY[address -> instr] =
                          , 2
                                 , 0),
  loader(0, (: (MOVE
                (MOVE
                          , 3
                                 , 0),
                (MOVE
                          , 4
                                 , 1),
                          , 4
                (SUB
                                 , 2),
                          , 4
                (JUMPZ
                                 , 12),
                          , 2
                                , χ(0)),
                (INCR
                (MOVERIND, 4
                                , 2),
                (MOVERIND, 5
                                , 3),
                (SUB
                          , 5
                                , 4),
                                , 2),
                          , 5
                (JUMPZ
                                , 2),
                          , 3
                (MOVE
                (JUMP
                          , 2
                                , X(1)),
                (RET
                          , X(2), X(3)) :))
```

If called with two addresses i and j in memory locations 0 and 1, the program MIN leaves the address of the minimum content of the array from i through j in memory location 2. Consider the following state S1.

```
STK: list[address]
MEM: [address -> nat]
S1: state =
           := 0,
 (# pc
           := STK,
    stk
           := MEM WITH
    mem
                              (1) := 15, (6) := 102, (7) := 111,
               [(0) := 6,
                (8) := 103, (9) := 103, (10) := 103, (11) := 101,
                (12) := 103, (13) := 103, (14) := 101, (15) := 103],
    halted := false,
    code
           := MIN
  #)
```

The stack component is uninterpreted and all memory locations except for a finite number are "don't cares". If we (symbolically) evaluate the stack machine interpreter sm.exec in 5 on the argument (100, S1) in PVS we get a list of 73 states. The first entry of this list represents the final state of the computation. Thus car(sm.exec(100, S1)) evaluates to:

In order to observe the dynamic behavior of the memory location mem(s)(3) one produces a new interpreter smo.exec by instantiating the generic interpreter exec in the following way.

Now, symbolic evaluation of smo.exec(100, S1) yields a list of the values of the result location during evaluation.

```
(: 11, \ldots, 11, 6, \ldots, 6, X(3) :)
```

Symbolic evaluation builds up huge expressions and soon becomes unmanageable when a large number of conditions evaluate neither to true nor false.

Simple evaluation of car(smo.exec(100, S2)), for example, essentially yields unfolded expression trees, since conditions like MEM2(12) < MEM2(6) or MEM2(14) = MEM2(6) can not be decided by evaluation but only by use of the axiom MEM2\_ax 10. The symbolic evaluator calls in these situations the PVS prover with the current context as hypotheses, the condition to be decided as the proof goal, and a suitably defined proof strategy; hereby, arbitrary strategies—including user-defined tactics—of the PVS prover can be used to decide conditions. For the example above, it suffices to apply a tactic that finds appropriate instances of the axiom MEM2\_ax followed by a call to the PVS decision procedures in order to simplify the expression car(smo.exec(100, S2)) to get the desired outcome 6.

#### 5 Conclusion

We have described a symbolic simulator for a functional subset of the PVS specification language and demonstrated its usage for validating a simple assembler program for a stack machine by applying it to incomplete data.

The main characteristics of our symbolic evaluator are its efficiency due to compilation of PVS functions to Lisp, retranslations of Lisp closures to PVS lambda-expressions, and the usage of theorem proving to decide conditions involving uninterpreted constants, uninterpreted function symbols, or even quantified boolean expressions.

The effectiveness of animating PVS specifications has been demonstrated through validation of a number of small to medium sized case studies. Besides the toy stack machine [21] described in this paper we have used animation of specifications to validate PVS models of microprocessors like the Transputer or the (pipelined) DLX [2, 19]. Further application of symbolic simulation include animation of the denotational semantics of imperative programs [14], validation of bisimulation diagrams, and partial evaluation of functions and state machines. Moreover, our evaluator has not only been used for the validation of formal models but has also shown to be useful in the context of theorem proving itself. A variant of our simulator, restricted to ground expressions, has recently been added to the main simplification strategy of the PVS prover [16]. Furthermore, efficient evaluation proved to be a necessity for safely extending theorem proving capabilities by replacing deduction with the evaluation of meta programs [20]

Transcriptions of abstract state machine (ASM) [6] models like the ones described in [2] into PVS demonstrate that symbolic evaluation can readily be used to animate deterministic ASMs. Basically, so-called *dynamic functions* are translated to state transformers and a centralized case analysis, like the one for stack machine above, is used to dispatch the ASM rules. Since the majority of ASM specifications we have encountered can easily be rewritten in this way, our symbolic simulator may form a viable alternative to specialized ASM simulators.

Animation of PVS specification has proven to be quite effective for our case studies. Although PVS is not built for effective symbolic evaluation and will run forever on larger examples, with our compilation methods it executes around 20

stack machine instructions per second on a SUN Ultra 1 in the naive implementation and around 100000 instructions per second with the destructive update technique. However, this may still not be efficient enough for animating myriads of test cases for state machines like industrial-sized microprocessors.

In order to further improve efficiency of symbolic evaluation, we plan to develop a compiler from PVS to LISP that uses the power of the PVS type system (including a state monad) in a systematic way to produce efficient code. Equally important, an interactive environment for the programming language PVS including the usual infrastructure like tracing, stack inspection, breakpointing, and statistical and diagnostic information is necessary for exploring expression values during execution.

Acknowledgements. The initial transcription of the stack machine to PVS has been performed by J. Rushby and M. Wilding sent us his modified specifications. We also thank H. Schwichtenberg for providing us with his Scheme programs for implementing inverse evaluation.

### References

- U. Berger and H. Schwichtenberg. An inverse of the evaluation functional for typed λ-calculus. In Proceedings, Sixth Annual IEEE Symposium on Logic in Computer Science, pages 203-211, Amsterdam, The Netherlands, 15-18 July 1991. IEEE Computer Society Press.
- E. Börger and S. Mazzanti. A Correctness Proof for Pipelining in RISC Architectures. Technical Report DIMACS 96-22, Dipartimento di Informatica, University of Pisa, Corso Italia 40, 56125 Pisa, Italy, 1996.
- R.S. Boyer and J.S. Moore. Mechanized Formal Reasoning about Programs and Computing Machines. In Automated Reasoning and Its Applications: Essays in Honor of Larry Wos. MIT Press, 1996.
- G. Dowek, A. Felty, H. Herbelin, G. Huet, Ch. Murthy, C. Parent, Chr. Paulin-Mohring, and B. Werner. The Coq Proof Assistant User's Guide (Version 5.8). INRIA-Rocquencourt - CNRS - ENS Lyon. Projet Formel.
- M.J.C Gordon and T.F. Melham. Introduction to HOL: A Theorem Proving Environment for Higher-Order Logic. Cambridge University Press, 1993.
- Y. Gurevich. Evolving algebras 1993: Lipari guide. In E. Börger, editor, Current Trends in Theoretical Computer Science, pages 9-36. Computer Science Press, 1995.
- D. Hardin, M. Wilding, and D. Greve. Transforming the Theorem Prover into a Digital Design Tool: From Concept Car to Off-Road Vehicle. In CAV'98: Computer-Aided Verification, Lecture Notes in Computer Science. Springer Verlag, June 1998.
- S.L. Peyton Jones and Ph. Wadler. Imperative Functional Programming. In ACM Symposium on Principles of Programming Languages (POPL'93), pages 71-84, Charleston, January 1993.
- M. Kaufmann and J S. Moore. An Industrial Strength Theorem Prover for a Logic based on Common Lisp. *IEEE Transactions on Software Engineering*, 21(4):203– 213, 1997.

- J S. Moore. Symbolic Simulation: an ACL2 Approach. In Formal Methods in Computer-Aided Design (FMCAD '98), Lecture Notes in Computer Science. Springer Verlag, 1998. Accepted for publication.
- 11. S. Owre, J. Rushby, N. Shankar, and F. von Henke. Formal Verification for Fault-Tolerant Architectures: Prolegomena to the Design of PVS. *IEEE Transactions on Software Engineering*, 21(2):107-125, February 1995.
- S. Owre, N. Shankar, and J.M. Rushby. The PVS Specification Language. Computer Science Lab, SRI International, Menlo Park CA 94025, March 1993.
- 13. L.C. Paulson. *Isabelle: A Generic Theorem Prover*. Number 828 in Lecture Notes in Computer Science. Springer-Verlag, 1994.
- H. Pfeifer, A. Dold, F. W. v. Henke, and H. Rueß. Mechanized Semantics of Simple Imperative Programming Constructs. Ulmer Informatik-Berichte 96-11, Universität Ulm, Fakultät für Informatik, 1996.
- V. Pratt. Anatomy of the Pentium Bug. In P.D. Mosses, M. Nielsen, and M.I. Schwartzbach, editors, TAPSOFT'95: Theory and Practice of Software Development, number 915 in Lecture Notes in Computer Science, pages 97-107. Springer Verlag, May 1995.
- 16. N. Shankar. Personal communication. 1998.
- M. Srivas, H. Rueß, and D. Cyrluk. Hardware verification using pvs. In Th. Kropf, editor, Formal Hardware Verification Methods and Systems in Comparison, volume 1287 of Lecture Notes in Computer Science, chapter 4, pages 156-205. Springer Verlag, 1997.
- M.K. Srivas and S.P. Miller. Formal Verification of the AAMP5 Microprocessor. In M.G. Hinchey and J.P. Bowen, editors, Applications of Formal Methods, International Series in Computer Science, chapter 7, pages 125–180. Prentice Hall, Hemel Hempstead, UK, 1995.
- 19. M. Stegmüller. Formale Verifikation des DLX RISC-Prozessors: Eine Fallstudie basierend auf abstrakten Zustandsmaschinen. Master's thesis, Universität Ulm, 1998. In German. Available at:
  - http://www.informatik.uni-ulm.de/ki/Edu/Diplomarbeiten.
- 20. Friedrich W. von Henke, Stephan Pfab, Holger Pfeifer, and Harald Rueß. Case Studies in Meta-Level Theorem Proving. In Jim Grundy and Malcolm Newey, editors, Proc. Intl. Conf. on Theorem Proving in Higher Order Logics, Lecture Notes in Computer Science. Springer-Verlag, September 1998. To appear.
- 21. M. Wilding. Robust Computer System Proofs in PVS. Presented at LFM'97: the Fourth NASA Langley Formal Methods Workshop; also available from http://www.csl.sri.com/sri-csl-fm.html, 1997.