Using Model Checking with Symbolic Execution to Verify Parallel Numerical Programs

Stephen F. Siegel
Dept. of Computer Science
University of Massachusetts
Amherst, MA 01003
siegel@cs.umass.edu

Anastasia Mironova
School of Computing
University of Utah
Salt Lake City, UT 84112
mironova@sci.utah.edu

George S. Avrunin and Lori A. Clarke
Dept. of Computer Science
University of Massachusetts
Amherst, MA 01003
{avrunin,clarke}@cs.umass.edu

ABSTRACT

We present a method to verify the correctness of parallel programs that perform complex numerical computations, including computations involving floating-point arithmetic. The method requires that a sequential version of the program be provided, to serve as the specification for the parallel one. The key idea is to use model checking, together with symbolic execution, to establish the equivalence of the two programs.

1. INTRODUCTION

In domains that require extensive computation, such as high-performance scientific computing, a program may be divided up among several processors working in parallel in order to reduce the overall execution time. The process of "parallelizing" a sequential program is notoriously difficult and error-prone. Attempts to automate this process have met with only limited success, and thus most parallel code is still written by hand. The developers of such programs expend an enormous amount of effort in testing, debugging, and a variety of ad hoc methods to convince themselves that their code is correct. Hence any techniques that can help establish the correctness of these programs or find bugs in them would be very useful.

In this paper we focus on parallel *numerical* programs, i.e., parallel programs that take as input a vector of (usually floating-point) numbers and produce as output another such vector. Examples include programs that implement matrix algorithms, simulate physical phenomena, or model the evolution of a system of differential equations. We are interested in techniques that can establish the correctness of a program of this type—i.e., prove that the program always produces the correct output for any input—or that exhibit appropriate counterexamples if the program is not correct.

The usual method for accomplishing this—testing—has two significant drawbacks. In the first place, it is usually infeasible to test more than a tiny fraction of the inputs that a parallel numerical program will encounter in use. Thus, testing can reveal bugs, but, as is well-known, it cannot show that the program behaves correctly on the inputs that are not tested. Secondly, the behavior of concurrent programs, including most parallel numerical programs, typically depends on the order in which events occur in different processes. This order depends in turn on the load on the processors, the latency of the communication network, and other such factors. A parallel numerical program may thus behave differently on different executions with the same input vector, so getting the correct result on a test execution does not even guarantee that the program will behave correctly on another execution with the same input.

The method proposed here, which combines model checking with symbolic execution in a novel way, does not exhibit these two limitations: it can be used to show that a parallel numerical program produces the right result on any input vector, regardless of the particular way in which the events from the concurrent processes are interleaved.

In attempting to apply model checking techniques in this setting, two issues immediately present themselves. The first arises from the fact that these techniques require that one first build a finite-state model of the program being checked. But numerical programs typically deal with huge amounts of floating-point data, and the very nature of our problem dictates that we cannot just abstract this data away. Hence it is not obvious how to construct appropriate finite-state models of the programs. The second issue concerns the nature of the property we wish to check: the statement that the output produced by the program is correct must be made precise, and formulated in some way that is amenable to model checking tools.

We deal with the first issue by modeling computations in the programs *symbolically*. That is, in our model, the input is considered to be a vector of symbolic constants x_i, and the output is some vector of symbolic expressions in the x_i. The numerical operations in the program are replaced by appropriate symbolic operations in the model. Furthermore, each symbolic expression is represented by a single integer, which prevents the blowup of the size of the state vector and which makes it possible to easily express the model in the language of standard model checking tools, such as Spin [9].

We deal with the second issue by requiring that the user provide a sequential version of the program to be verified, which will serve as a specification for the parallel one. The model checker will be used to show that the parallel and sequential programs are *equivalent*, i.e., that they produce the same output on any given input. Of course, this means that our method only reduces the problem of producing a correct parallel program to the problem of producing a correct sequential one. However, most problems in this domain have a much simpler sequential solution than parallel one, and it is already common for developers of scientific software to produce sequential versions of their parallel programs, for testing and other purposes. Moreover, we will see below that our method provides additional information that

can help verify the correctness of the sequential program as well.

Another issue that arises in this approach is the fact that most numerical programs contain branches on conditions that involve the input. Such programs may be thought of as producing a set of cases, each case consisting of a predicate on the input and the corresponding symbolic output vector. Our method deals with this as follows. We use the model checker to explore all possible paths of the sequential program, and for each such path we record the *path condition pc*, the Boolean-valued symbolic expression on the input that must hold in order for that path to have been followed. The model of the parallel program is engineered to take as input not only the symbolic input vector, but the path condition *pc* as well. The model checker is then used to explore all possible paths of the parallel program that are consistent with *pc*. If, for every *pc*, the result produced by the parallel program always agrees with the result produced by the sequential one, the two programs must be equivalent.

The method is described in detail in Section 2. Using SPIN, we have applied the method to four parallel numerical programs; we describe this experience and present some data that arose from it in Section 3. Section 4 discusses related work and Section 5 presents some conclusions and directions for future work.

2. METHODOLOGY

We consider a parallel numerical program P_{par} that consists of a fixed number of parallel processes. We write n for the number of parallel processes. We assume that these processes have no shared memory and communicate only through message-passing functions such as those provided by the *Message Passing Interface* (MPI) [15, 16]. (Though much of what follows will apply equally to other communication systems, or even to shared memory systems, MPI has become the *de facto* standard for high performance computation, particularly in the domain of scientific computation.) We assume we are given a sequential program P_{seq}, which serves as the specification for P_{par}. We also assume that both P_{seq} and P_{par} terminate normally on every input, a property that can often be verified using more traditional model checking techniques [21, 22]. In some cases, we may also have to impose a small upper bound on the number of iterations of certain loops in a program, to ensure that the model we build will not have an inordinately large (or even infinite) number of states.

Notice that the requirement that P_{par} and P_{seq} be equivalent implies, in particular, that each program be *deterministic*, i.e., that if given the same input twice, it will produce the same output. If either program fails to be deterministic, this will be caught and flagged as an error by our method.

To simplify the presentation, we begin by explaining the method under the assumption that neither program contains branches on expressions involving variables that are modeled symbolically. After this we consider some numerical issues that arise from the fact that floating-point arithmetic is only an approximation to the arithmetic of the real numbers, and finally we describe the general approach, in which branches on symbolically modeled expressions are allowed.

2.1 A simple example

To illustrate the method, we consider the example of Figure 1(a). This sequential C code takes the product of an $N \times L$ matrix A and an $L \times M$ matrix B and stores the result in the $N \times M$ matrix C. We can consider this to be a numerical program for which the input vector consists of the $NL + LM$ entries for A and B, and the output vector consists of the NM entries of C at termination. There are many ways to parallelize P_{seq}, but we will consider the one shown in Figure 1(b), which is adapted from [7] and uses MPI functions for interprocess communication. Each process should be thought of as executing its own copy of this code, in its own local memory. A process may also obtain its *rank* (a unique integer between 0 and $n-1$) from the MPI infrastructure. For this code, which uses a *master-slave* approach to achieve automatic load-balancing, we assume that $N \geq n-1 \geq 1$, and that all three matrices are stored in the local memory of the process of rank 0 (the *master*). To compute the product, the master will distribute the work among the processes of positive rank (the *slaves*).

We assume that each slave process already has a copy of B in its local memory. The master begins by sending the first row of A to the first slave, the second row of A to the second slave, and so on, until the first $n-1$ rows of A have been handed out. A slave, after receiving a row vector of length L from the master, multiplies it by B, and sends back the resulting row vector of length M to the master. The master waits at a receive statement that will accept a message from any process (we will refer to a statement of this kind as a *wildcard receive*). After one or more messages have arrived, the master chooses one for reception, copies the row vector received into the appropriate row in C, sends the next row of A to the slave that had just returned the result, and returns to the wildcard receive. It continues in this way until all the rows of A have been handed out. After that point, whenever a slave sends in a result, the master sends back a termination message to that slave. After all results have come in, and the last termination message has been sent out, C should contain the product of A and B, and all processes should terminate normally.

The first step of our method is to create a finite-state model M_{seq} of P_{seq} in Promela, the input language for SPIN. The model will use symbolic expressions in place of the floating-point values that arise in P_{seq}. (Integer values can also be modeled symbolically, though this is often not necessary.) A symbolic expression may be thought of as a tree-like structure in which the leaf nodes are either floating-point literals or symbolic constants. The symbolic constants are denoted $x_0, x_1, \ldots$ and correspond to the components of the input vector. To each non-leaf node in the tree is associated a (unary or binary) operator, e.g., $+,-,*,/$, or any other arithmetic operator that occurs in the program.

Each numerical operation in the program involving a symbolically modeled variable is replaced by an operation on symbolic expressions in the model. The symbolic operation simply forms a new tree from a given operator and one or two operands. We will use the usual infix notation to denote symbolic expressions, but we must keep in mind that no interpretation is given to the operations, and none of the usual rules of real arithmetic (associativity, commutativity, etc.) hold. For example, for the matrix multiplication program with $N = L = M = 2$, if the initial symbolic values for A and B are given by

$$A = \begin{pmatrix} x_0 & x_1 \\ x_2 & x_3 \end{pmatrix}, \quad B = \begin{pmatrix} x_4 & x_5 \\ x_6 & x_7 \end{pmatrix},$$

```
double A[N][L], B[L][M], C[N][M];

                ⋮

int i,j,k;
for (i=0; i<N; i++)
  for (j=0; j<M; j++) {
    C[i][j] = 0.0;
    for (k=0; k<L; k++)
      C[i][j] += A[i][k]*B[k][j];
  }
```

(a) Sequential code

```
int rank,nprocs,i,j,numsent,sender,row,anstype;
double buffer[L], ans[M];
MPI_Status status;
MPI_Comm_size(MPI_COMM_WORLD, &nprocs);
MPI_Comm_rank(MPI_COMM_WORLD, &rank);
if (rank==0) { /* I am the master */
  numsent=0;
  for (i=0; i<nprocs-1; i++) {
    for (j=0; j<L; j++)
      buffer[j] = A[i][j];
    MPI_Send(buffer, L, MPI_DOUBLE, i+1,
      i+1, MPI_COMM_WORLD);
    numsent++;
  }
  for (i=0; i<N; i++) {
    MPI_Recv(ans, M, MPI_DOUBLE, MPI_ANY_SOURCE,
      MPI_ANY_TAG, MPI_COMM_WORLD, &status);
    sender = status.MPI_SOURCE;
    anstype = status.MPI_TAG-1;
    for (j=0; j<M; j++)
      C[anstype][j] = ans[j];
    if (numsent<N) {
      for (j=0; j<L; j++)
        buffer[j] = A[numsent][j];
      MPI_Send(buffer, L, MPI_DOUBLE, sender,
        numsent+1, MPI_COMM_WORLD);
      numsent++;
    }
    else MPI_Send(buffer, 1, MPI_DOUBLE, sender,
      0, MPI_COMM_WORLD);
  }
} else { /* I am a slave */
  while (1) {
    MPI_Recv(buffer, L, MPI_DOUBLE, 0,
      MPI_ANY_TAG, MPI_COMM_WORLD, &status);
    if (status.MPI_TAG==0) break;
    row = status.MPI_TAG-1;
    for (i=0; i<M; i++) {
      ans[i] = 0.0;
      for (j=0; j<L; j++)
        ans[i] += buffer[j]*B[j][i];
    }
    MPI_Send(ans, M, MPI_DOUBLE, 0,
      row+1, MPI_COMM_WORLD);
  }
}
```

(b) Parallel code

Figure 1: Matrix multiplication code excerpts

then the final value of $C[0][0]$ will be the symbolic expression $(0.0 + x_0x_4) + x_1x_6$, which does not equal the symbolic expression $x_0x_4 + x_1x_6$. The symbolic structure can be represented in a language such as Promela using standard data structures such as integer arrays, but we will see shortly that this is not really necessary.

The next step of our method is to create a finite-state model M_{par} of P_{par}. To do this we use Spin processes to represent the processes of the parallel program and Spin channels to transfer messages between processes, using techniques such as those of [22] and [23]. The arithmetic operations are represented symbolically, just as in the sequential case. Finally, a *composite model* is formed, in which first M_{seq} is executed in its own Spin process, then M_{par} is executed using n additional Spin processes, and finally a series of assertions are checked to verify that the final symbolic entries of the copy of C generated by M_{seq} agree with those generated by M_{par}. Spin is then used to explore all possible paths of the composite model and to verify that the assertions are never violated. In the matrix multiplication example, there are many such paths, due to all the different possible orders in which the slaves can return their results to the master.

Now, the method described above may work for small models, but it has a serious drawback. For a typical program, the size of the symbolic expressions—and therefore the size of the structure used to represent the state of the model—can quickly blow up. Like most model checking tools, Spin stores the set of states it has encountered as it searches the state space of the model, and the amount of memory required to represent this set is usually the main barrier to a successful completion of the search. Since the memory required to represent the set is approximately the product of the number of states and the size of the structure used to represent a single state, the method we have proposed has little chance of scaling.

To ameliorate this problem, we use a form of *value numbering* to reduce the memory needed to represent a symbolic expression and use *subexpression sharing* to reduce the total number of expressions and facilitate expression comparison. Using this approach, the floating-point values in the original programs are represented by integer indices that refer to entries in a static *symbolic expression table*. (By *static*, we do not mean that the table never changes, but that it is shared by every state in the state space, just as a static variable in a Java class is shared by all instances of that class.) The table contains one entry for every expression (including every subexpression of every expression) that is encountered during the search of the state space of the composite model. An entry for a binary expression is a triple in which the first component is an operator code, the second component is an integer referring to an (earlier) entry in the table corresponding to the left operand, and the third component is an integer corresponding similarly to the right operand. The entry for a unary expression is similar but has only two components. An entry for a leaf expression has either the form $(\texttt{X}, i)$, corresponding to the symbolic constant x_i, or $(\texttt{L}, \alpha)$, where α is a floating-point number, corresponding to a literal value.

The table is initialized by entering the literal values 0 and 1, as these are needed by many models and by many of the routines in our symbolic manipulation package. Next, the symbolic constants for the input vector are entered into the

i	e_i	interpretation
0	$(\texttt{L}, 0.0)$	0.0
1	$(\texttt{L}, 1.0)$	1.0
2	$(\texttt{X}, 0)$	x_0
3	$(\texttt{X}, 1)$	x_1
⋮	⋮	⋮
9	$(\texttt{X}, 7)$	x_7
10	$(*, 2, 6)$	x_0x_4
11	$(+, 0, 10)$	$0.0 + x_0x_4$
12	$(*, 3, 8)$	x_1x_6
13	$(+, 11, 12)$	$(0.0 + x_0x_4) + x_1x_6$
14	$(*, 2, 7)$	x_0x_5
⋮	⋮	⋮
25	$(+, 23, 24)$	$(0.0 + x_2x_5) + x_3x_7$

Figure 2: Symbolic expression table for 2×2 matrix multiplication

table. Other entries to the table are made as needed as symbolic operations are performed during the search of the state space. The arithmetic operations are modeled by operations on integers that refer to entries in the table. The operation that performs addition, for example, takes two integers i and j, and first looks in the table to see if the triple $(+, i, j)$ has already been entered. If it has, the addition operation returns the index of that triple. If it has not, it appends that triple to the end of the table and returns the new index. This guarantees that every expression has a unique entry in the table, and so the expressions corresponding to two integers i and j are equal if and only if $i = j$.

The table that is constructed during the verification of the 2×2 matrix multiplication example is excerpted in Figure 2. At the end of execution of M_{seq}, the table will have 26 entries. Along the way, $C[0][0]$ takes on the values 0, then 11, and finally 13. Hence one of the assertions that will be checked is that, at the termination of any execution of M_{par}, the variable $C[0][0]$ in the master process will also be 13.

In this case, when the state space of M_{par} is explored, no new entries are ever made, because all of the expressions generated can already be found in the table. (In more complicated examples, however, the parallel program may also add new expressions.) In fact, for non-trivial sizes (see Section 3), Spin can verify that the assertions are never violated, establishing the equivalence of the two programs.

2.2 Numerical Issues

Floating-point arithmetic is only an approximation to the arithmetic of real numbers, and many of the standard properties of the latter do not necessarily hold for the former [5]. (The exact differences depend on which particular floating-point arithmetic one uses.) In the matrix multiplication example, the symbolic expressions computed by the sequential and parallel models are exactly the same, which guarantees that the programs being modeled will always produce the same results, no matter what arithmetic is used to execute the programs (assuming, of course, that the arithmetic functions are deterministic). There are cases, however, where two models may compute expressions that are not exactly the same, but which may be close enough for particular needs. For example, in most floating-point arithmetics—including all those that conform to the IEEE 754 or 854 standards [10, 11]—the expressions $0 + f$ and f must always evaluate to the same floating-point value, for any floating-point expression f. Hence, if the symbolic results produced by the two models are the same "up to" the relation that identifies any symbolic expression e with the symbolic expression $0+e$, we are still guaranteed that the two programs will produce the exact same floating-point results on any platform implementing IEEE arithmetic.

In general, let $\sim$ be an equivalence relation on the set $S(X)$ of symbolic expressions over a set of symbolic constants $X = \{x_1, x_2, \ldots\}$. We assume that $\sim$ is *operation-preserving*, i.e., that

$$e_1 \sim e_2 \wedge f_1 \sim f_2 \;\Rightarrow\; e_1 + f_1 \sim e_2 + f_2$$

holds for all $e_i, f_i \in S(X)$, and that similar statements hold for the other operators. This means that each operation induces an operation on the set of equivalence classes $\bar{S}(X) \equiv S(X)/\sim$, and so all of the arithmetic and comparisons for equality in the models may be thought of as taking place in $\bar{S}(X)$.

Note that in $\bar{S}(X)$, it is no longer trivial to test for the equality of two elements. We will see in Section 3.1 that our implementation deals with this by performing certain simplifications on an expression before it is entered into the symbolic table. This is not quite as strong as reducing the expression to a true normal form (i.e., to a unique representative of its equivalence class), but it is very inexpensive and provides sufficient precision for most cases.

Each operation-preserving equivalence relation yields a different notion of program equivalence. We have identified three that we think are useful and have used in our implementation, though the same methods can certainly be used for other relations. The three relations are as follows:

- *Herbrand equivalence*: this is the strongest, and therefore most desirable, notion of equivalence. Two symbolic expressions are Herbrand equivalent if and only if they are exactly equal. As we have seen, two Herbrand equivalent programs will produce the same results, independently of the way in which the arithmetic operations are implemented.

- *IEEE equivalence*: this is a slightly weaker relation. There are a number of identities for real arithmetic that also hold for IEEE arithmetic, e.g, $x + y = y + x$, $xy = yx$, and $1x = x1 = x+0 = 0+x = x/1 = x$. Two elements of $S(X)$ are considered to be equivalent if one can be transformed to the other by a finite sequence of transformations corresponding to such identities. Two IEEE equivalent programs must produce the same output on any platform implementing IEEE arithmetic. Of course, they would also produce the same output if the arithmetic were exactly real arithmetic.

- *Real equivalence*: this is weaker still. Two elements of $S(X)$ are considered to be equivalent if one can be transformed to the other using any identities of real numbers, including those that do not hold for IEEE arithmetic, such as the associativity of addition or multiplication, and the distributive property. Two real equivalent programs would produce the same results if all computations were performed as real arithmetic, but they may produce different results when run on

an actual computer, even one that implements IEEE arithmetic. The differences may be slight, but in some situations the error can mushroom and the two can differ greatly.

The sad truth is that real equivalence is often the best that we can hope for. This is because there are many common scenarios that rely on associativity or some other property that does not hold for IEEE arithmetic. For example, it is often the case that one needs to compute a sum of floating-point variables that reside in the local memory of different processes and return the result to every process. MPI provides a convenient way to do this: one just calls `MPI_Allreduce` with a parameter specifying that the *reduction operation* is to be floating-point addition. However, the MPI Standard states that the implementation may add the values in any order—the implementation is not even required to use the same order twice. Hence an MPI program making one call to `MPI_Allreduce` may produce different results when run twice on the same input, even if the execution platform uses IEEE arithmetic.

For programs that are real but not IEEE equivalent, difficult issues may arise in creating test oracles or in determining whether the error (the difference between the actual results and what the results would have been had real arithmetic been used) falls within acceptable bounds. Such questions are simply beyond the scope of our method. Other investigations have attempted to deal precisely with floating-point errors, in different circumstances; see, for example, [14] and the references cited there.

We mentioned above that sometimes we might want to model integer variables symbolically. This requires a small modification to the above framework, in which we associate a type (either integer or floating-point) to each symbolic constant and, consequently, a type to each symbolic expression. The notion of Herbrand equivalence is unchanged, but for both IEEE and real equivalence we allow all the usual rules of integer arithmetic, including commutativity, associativity, and the distributive property for integer addition and multiplication.

2.3 The general case

The method used for the matrix multiplication example applies to any program with no branches on expressions that involve the symbolically modeled variables. We now drop this restriction. To illustrate the general case, we use the program in Figure 3, which implements the Gaussian elimination algorithm to transform an $N \times M$ matrix to its reduced row echelon form. The input vector for this program consists of the NM initial values of the matrix entries, and the output vector consists of the NM final values of those entries.

Recall that an important step in this algorithm is to locate, at each stage, a *pivot row*, i.e., a row at or below the current top row that contains a non-zero entry in the current column. This is accomplished in the sequential code by looping over the rows, starting at `top` and working down, looking for a non-zero entry. If none is found, the algorithm moves to the next column and loops over the rows again. This continues until the first non-zero entry is found, or until we fall off the bottom or the right side of the matrix.

In the parallel version (Appendix B.1) , we assume that $n = N$ (where n is the number of parallel processes) and that the i^{th} row of the matrix is stored in the local memory

```
double matrix[N][M];

                ⋮

int top,col,row,i,j;
double pivot,tmp;
for (top=col=0; top<N && col<M; top++, col++) {
  pivot = 0.0;
  for (; col<M; col++) {
    for (row=top; row<N; row++) {
      pivot = matrix[row][col];
      if (pivot!=0.0) break;
    }
    if (pivot!=0.0) break;
  }
  if (col>=M) break;
  if (row!=top)
    for (j=0; j<M; j++) {
      tmp = matrix[top][j];
      matrix[top][j] = matrix[row][j];
      matrix[row][j] = tmp;
    }
  for (j=col; j<M; j++) matrix[top][j] /= pivot;
  for (i=0; i<N; i++)
    if (i!=top) {
      tmp = matrix[i][col];
      for (j=col; j<M; j++)
        matrix[i][j] -= matrix[top][j]*tmp;
    }
}
```

Figure 3: Sequential Gaussian elimination code

of the process of rank i. The pivot row is determined in a very different way, using a call to `MPI_Allreduce` in which the reduction operation returns the minimum of the given values. Each process contributes an integer to this communication, according to the following rule: if its entry in position `col` is 0 or the rank of the process is less than `top`, the process contributes the integer n, else it contributes its rank. The call to `MPI_Allreduce` results in the minimum of all these contributions being stored in the variable `row` of each process. If, after this communication completes, `row` is less than n, then each process knows that the process of rank `row` will be used as the next pivot row and breaks out of the pivot-searching loop, else the search for a pivot continues. Additional communication is used to exchange the top and pivot rows and to broadcast the pivot row.

Because of the branch expressions that involve the floating-point input (e.g., `pivot!=0.0`), the sequential program can follow different paths, depending on the input. Consider, for example, the case where $N = M = 2$, and the matrix is initially $\begin{pmatrix} x_0 & x_1 \\ x_2 & x_3 \end{pmatrix}$. If $x_0 \neq 0$ and $x_3 - x_2(x_1/x_0) = 0$ then the program will follow a path resulting in the final value of $\begin{pmatrix} 1 & x_1/x_0 \\ 0 & 0 \end{pmatrix}$ (assuming IEEE arithmetic is used). If instead $x_0 \neq 0$ and $x_3 - x_2(x_1/x_0) \neq 0$, the final result is the identity matrix. In fact, in this 2×2 case, there are 7 possible paths through the sequential program. To each path there is an associated *path condition*, the predicate on the input vector that must hold in order for that path to be

followed, and a resulting symbolic output vector. (Notice it is possible for two different paths to yield the same output: the path arising from the condition $x_0 = 0 \wedge x_2 \neq 0 \wedge x_1 \neq 0$ also yields the identity matrix.)

We deal with this as follows. In M_{seq}, we model the floating-point variables symbolically, as before, but we also introduce an integer variable that gives the index in the symbolic table of the current path condition *pc* for the program. This expression is Boolean-valued and can involve operators such as $<,>,=,\neq,\geq,\leq,\wedge,\vee$. Its initial value is the special symbolic expression **true**. At each point where there is a floating-point branch in the program, say on a condition e, the model calls a function $\phi(pc, e)$. This function returns one of three possible values: if it can determine that $pc \Rightarrow e$ it returns **true**; if it can determine that $pc \Rightarrow \neg e$, it returns **false**; and if it cannot determine either, it returns **unknown**. If the answer is **true** or **false**, the corresponding branch is taken, but if the answer is **unknown** then the model makes a non-deterministic choice between the true and false branches. In this latter case, if the true branch is selected, the value of *pc* is updated by setting it to $pc \wedge e$, while if the false branch is selected, it is set to $pc \wedge \neg e$.

Recall that in the composite model, the execution of M_{par} begins just after M_{seq} terminates. Now, the branches in M_{par} will be dealt with in the same way as in M_{seq}, but—and this is the crucial point—M_{par} will use the same path condition variable that was used in M_{seq}. Hence execution of M_{par} begins with *pc* holding the final value computed by the sequential model. This means that, *assuming $\phi(pc, e)$ can be evaluated with sufficient precision*, the parallel model can only follow a path that is consistent with the one followed by the sequential model. Finally, the last step in the composite model is the sequence of assertions comparing the output vectors of the two models, just as before.

Now when SPIN is used to check for assertion violations in the composite model, it will have to explore all possible paths through M_{seq}, and for each of these it will have determined a path condition-output vector pair $(pc, \mathbf{y})$. For each such pair, it will explore all possible paths of the parallel model that are consistent with *pc*, determine the parallel output $\mathbf{y}'$, and check the equivalence of $\mathbf{y}$ and $\mathbf{y}'$. If the assertions can never be violated, we can conclude that for any input vector, the two programs must produce equivalent results, assuming the arithmetic used in executing the programs obeys the identities of the designated equivalence relation.

Notice that the path condition *pc* produced by a sequential run does not necessarily specify all the branch conditions for M_{par}. In the Gaussian elimination code, for example, the sequential program breaks out of the loop that searches for a pivot as soon as the first non-zero entry is found. Hence the symbolic variables for the entries that are not examined remained unconstrained in *pc*. In the parallel code, on the other hand, each process examines its own entry to see if it is non-zero, and those processes that cannot make this determination based on *pc* must make a non-deterministic choice. The model checker explores all of these choices, and checks that each of them results in the same output vector $\mathbf{y}$.

The effectiveness of this approach depends heavily on the precision with which the $\phi(pc, e)$ are evaluated. If **unknown** is returned for a case where it can in fact be shown that $pc \Rightarrow e$ (or that $pc \Rightarrow \neg e$), it is possible that SPIN will explore *infeasible* paths through M_{seq}, or paths through M_{par} that are not consistent with the one followed by M_{seq}. In these cases the analysis might produce a spurious result, i.e., it might report that a violation has been found when one does not really exist. However, since the analysis is *conservative*, i.e., it only ignores a branch when it is certain that the branch cannot be taken, a positive result guarantees that the two programs are equivalent.

A useful byproduct of this method is the set of pairs $(pc, \mathbf{y})$ produced by SPIN in analyzing M_{seq}. These can be used to establish the correctness of the sequential program, although exactly how this is done would depend on the particular program. For the Gaussian elimination example, each of the matrices that corresponds to a $\mathbf{y}$ in one of the pairs can be checked, by a series of assertions, to satisfy the conditions of reduced row echelon form.

In summary, our method to compare a sequential and parallel version of a numerical program consists of the following steps: (1) build a SPIN model M_{seq} of the sequential program in which the floating-point computations (and perhaps some integer computations) are represented symbolically, and in which branches are modeled using non-deterministic choices and a path condition variable; (2) in a similar way, build a SPIN model M_{par} of the parallel program; (3) put these together to form a composite model, in which first M_{seq} is executed, then M_{par} (using the same path condition variable), and which ends with assertions stating that the outputs of M_{seq} and M_{par} agree; (4) use SPIN to check that the assertions of the composite model can never be violated.

3. EXPERIMENTS

In this section, we discuss our implementation of the method and our experience in applying it to four numerical programs. The source code for the implementation, the models, and all of the experimental results can be obtained at `http://laser.cs.umass.edu/~siegel/projects`.

3.1 Implementation

The core of our implementation is a library of functions for manipulating symbolic expressions and maintaining the symbolic expression table. This library is written in C and is incorporated into our models by using the embedded C code facility of SPIN. The entries in the table are C `structs`, and include fields for the index of the expression, an integer code representing the operator, pointers to the left and right subexpressions for binary expressions, etc. A hashtable is also used, in order to find table entries quickly.

Three "levels" of each arithmetic operation are provided, corresponding to the three different equivalence relations discussed in Section 2.2. The level 0 operations correspond to Herbrand equivalence; these simply form a new expression from the operator and operands, check to see if the new expression already exists in the table, add it to the table if it does not, and return the index. The level 1 operations correspond to IEEE equivalence and do a little more work; the level 1 addition operation, for example, checks to see whether one of the operands is 0 (in which case it returns the other operand), or if one operand is the negative of the other (in which case it returns 0). The level 2 operations correspond to real equivalence. The level 2 addition operation, for example, exploits associativity and commutativity to reduce sums to a simplified form in which the parentheses are moved to the left as far as possible, the terms are or-

dered by increasing index, literal integer terms are combined into a single literal, and so on. Of course, when checking for IEEE equivalence, the level 2 addition and multiplication operations can also be used for all integer expressions.

Similar things could of course be done for the other level 2 operations (multiplication, subtraction, and so on), but in the examples we have studied so far this has not been necessary, and so at present these work exactly as the corresponding level 1 operations. In our experience, it seems that additive reduction operations, which are very common in parallel numerical programs, account for much of the difference between the exact symbolic expressions computed in the sequential and parallel models. Reductions over other operations, such as multiplication, seem to be much less common. In any case, the symbolic package is designed to make it easy to specify the symbolic operation used for a particular computation in the code and to add new versions of the symbolic operations to reduce expressions to other simplified forms, as the need arises.

The function ϕ, which attempts to determine whether the given path condition pc implies the given expression e (or $\neg e$), is implemented as follows. First, by construction, pc will always be a conjunction of smaller expressions of the form pc_i or $\neg pc_i$, where each pc_i arises by evaluating one of the conditional expressions in a branching statement. Our implementation of ϕ simply loops over i, looking for a pc_i which can be easily seen to imply e or $\neg e$. By "easily seen" we mean by using reasoning such as $x < y \Rightarrow x \neq y$, $x = y \Rightarrow \neg(x \neq y)$, and so on. If it finds such a pc_i, it returns **true** or **false**, as the case may be, otherwise it returns **unknown**. This lightweight procedure seems to be effective because the conditional expressions evaluated in the sequential and parallel programs—or, at least, the ones that matter for correctness—tend to be quite similar.

All of the variables from the symbolic package are static, that is, they are not incorporated into the state vector by using, for example, the Spin `c_track` function. Thus the only variables incorporated into the state vector are those corresponding to variables in the original programs and the path condition variable.

The type of equivalence that one wishes to verify (Herbrand, IEEE, or real) is controlled by a command-line argument specified when compiling the verifier (pan.c) generated by Spin. This argument tells the symbolic package which level of symbolic operations it should use: level 0 for Herbrand, level 1 for IEEE, and level 2 for real. (When verifying IEEE equivalence, integer addition and multiplication operations use level 2, instead of level 1.) Finer control over the operations can be obtained by defining additional functions and calling them from the Promela model where desired.

3.2 The programs

For our preliminary study, we analyzed four scalable parallel numerical programs. We attempted to verify each of these using the method of this paper, scaling until Spin exhausted the 800 MB of available memory or verification time exceeded 10,000 seconds. In what follows, we give a brief description of each program, we discuss certain issues that arose in verifying its correctness, and we explain what we were able to verify (or not verify).

Partial orders and related techniques play an important role in model checking, by reducing the number of states that need to be explored. Ideally, we would have liked to apply techniques that are optimized for models of MPI programs, such as those discussed in [21], but we could not find an easy way to implement them in Spin. Instead, we proved a result (Appendix A.1) that justifies slightly weaker techniques, but can be easily incorporated into Spin models. One consequence of the theorem is that for models with no wildcard receives, we may instruct Spin to use only synchronous communication, and we may place the code for each process in an `atomic` block. Though this greatly restricts the ways in which events from the different processes can be interleaved, the theorem implies that Spin will still explore every possible terminal state of the model, which is all that is required for our purposes. For models with wildcard receives, we must use asynchronous communication, but we can still use `atomic` blocks, as long as every wildcard receive occurs either outside, or as the first statement of, an `atomic` block. In both cases, the reduction in the number of states explored can be dramatic.

In Table 1, we give data for the largest configuration of each program that we were able to verify. The columns of the table give (1) the type of equivalence that was verified, (2) the number n of parallel processes, (3) the number of distinct sequential executions, (4) the number of expressions generated in the course of the verification, (5) the length of the input vector, i.e., the number of symbolic constants, (6) the length of the output vector, (7) the maximum number of terms in the path condition conjunction, (8) the number of states explored, (9) the amount of memory used by Spin, and (10) the verification time. We used Spin version 4.2.4 with options `-DCOLLAPSE -DSAFETY -DNOBOUNDCHECK` on a 2.2GHz Pentium 4 Linux box.

3.2.1 *matmat*

Our first example is the matrix multiplication program of Section 2.1, with $N = L = M = 2(n-1)$. We were able to verify that the sequential and parallel programs are Herbrand equivalent for $n \leq 6$.

3.2.2 *gauss*

Our second example is the Gaussian elimination program of Section 2.3, with $N = M = n$. We were able to verify that the sequential and parallel programs are Herbrand equivalent for $n \leq 6$. We note, however, that in order to show that the sequential program really produces the reduced row echelon form, we needed IEEE arithmetic. This is because, for example, the use of Herbrand arithmetic results in matrix entries of the form x_0/x_0 where the definition of reduced row echelon form requires 1.

3.2.3 *jacobi*

Our third example implements a Jacobi iteration algorithm to solve a linear system of the form $A\mathbf{x} = \mathbf{b}$. Both the sequential and parallel versions are from the CD ROM accompanying [12]. In this algorithm, the $N \times N$ matrix A and the column vector $\mathbf{b}$ of length N form the input, and the goal is to solve for the value of the column vector $\mathbf{x}$ of length N, which forms the output. We take $N = 2n + 2$. The algorithm begins with an initial guess for $\mathbf{x}$ (the column vector in which every entry is 1.0), and then enters a loop in which the entries of $\mathbf{x}$ are adjusted at each iteration, based on the values of neighboring entries. The algorithm stops when the error term, which is computed as the inner product of the difference between two consecutive values of

program name	equivalence type	parallel processes	sequential executions	symbolic expressions	input vector	output vector	path condition	states (10^3)	memory (MB)	time (s)
matmat	Herbrand	6	1	2202	200	100	0	4443	217	506
gauss	Herbrand	6	13327	247656	36	36	36	16114	801	3224
jacobi	real	17	4	8239	1333	36	3	6295	362	9846
monte	IEEE	9	4	1232	99	1	3	3112	279	738

Table 1: Experimental data

x with itself, falls below a given threshold ϵ, or when the number of iterations exceeds a fixed bound `MAXITS`.

In the parallel version, the data are partitioned so that each process contains a certain number of rows of A, **x**, and **b**. Communication is used to update the contents of *ghost-cells*, which mirror the boundary data on neighboring processes, and in a reduction operation used to compute the error term after each iteration. In our models, ϵ is treated as another symbolic constant, and we take `MAXITS` $= 3$. (Some constant bound must be specified for `MAXITS` if the model is to have a finite number of states.)

Our analysis quickly revealed that the results of the sequential and parallel programs could disagree for $n = 2$, even using real arithmetic. The source of the problem was a small mistake in the computation of the error in the sequential code: instead of taking the inner product of the difference between two successive values of **x** with itself, the code simply took the inner product of the two successive values. This was pointed out to one of the authors, who acknowledged the mistake. After correcting this error, we verified real equivalence (which is the best that can be hoped for, due to the floating-point reduction operation) for $n \leq 17$. While this example scaled significantly further than the others, it is also the only case in which time, rather than memory, proved to be the limiting factor. This appears to be due to the large amount of computation required to simplify expressions when using the level 2 operations.

3.2.4 monte

Our fourth example is a parallel program taken from [7] that implements a Monte Carlo algorithm to estimate π. The algorithm repeatedly chooses a point at random from a square with sides of length 2. If the distance from the point to the center exceeds 1.0, an integer variable `out`, initially 0, is incremented, else a variable `in` is incremented. The estimate for π is `4.0*in/(in+out)`. The algorithm stops when an error term falls below a fixed threshold ϵ, or `in+out` exceeds a fixed bound. In the parallel code, one process acts as a random number server, returning blocks of random numbers to the remaining "worker" processes. The worker processes use these blocks to determine a set of points and make their own local `in` and `out` calculations. The values of `in` and `out` are summed at the end of each iteration, using an integer reduction operation. At the end of the reduction, each process has the global sums, forms the estimate for π, computes the error, and decides whether to perform another iteration or terminate.

The random nature of this code presents an interesting challenge to our method. On the face of it, a program that depends in an essential way on the values returned by a random function can hardly be deterministic. We resolve this problem by considering the sequence of random numbers generated by the random function to be the *inputs* to the program. Hence our method can be used to verify that if the random number function were to generate the same sequence of values for the sequential and the parallel programs, the two programs must return the same estimate for π. This seems to us to be a natural extension of our notion of equivalence to numerical programs that use random numbers.

We used two additional reduction techniques for this example, which proved very effective. The first concerns the program statement

```
if (x*x+y*y<1.0) then in++ else out++;
```

which is used to determine whether a point $(\mathtt{x}, \mathtt{y})$ is within distance 1.0 of the center. If we were to follow our method strictly, each time this statement is executed in M_{seq} a nondeterministic choice would be made between the two alternatives. Since this statement is executed many times in the model, the number of sequential executions would blow up quickly. To avoid this problem, we made a simple program transformation. First, we defined a new operation `delta` which takes two floating-point arguments a and b, and returns the integer 1 if $a < b$ and 0 otherwise. The statement above can then be replaced by

```
in += delta(x*x+y*y,1.0);
out += 1-delta(x*x+y*y,1.0);
```

which does not require a non-deterministic choice. The only change we had to make to the symbolic package was to add a level 0 operation for `delta`, i.e., we just treat `delta` as an uninterpreted function. With this modification, the symbolic output of M_{seq} will be a more complicated expression, involving many `delta`-subexpressions, but the number of executions of M_{seq} will be much smaller, which turns out to be a good tradeoff. Notice also what happens if we use IEEE arithmetic to compute the sum of `in` and `out`; since the symbolic package knows to use associativity and commutativity for integer expressions, the `delta` terms in the sum all cancel and the result is a single integer constant. This also reduces the number of states explored, since it allows the symbolic package to determine with precision when the sum exceeds the upper bound, rather than forcing it to make another non-deterministic choice.

The second reduction technique exploited a *symmetry reduction theorem* (Appendix A.2) that we proved for general parallel numerical programs. To see how this comes into play in this example, observe that in M_{par}, the worker processes can send their requests to the random number server in any order. Hence in one execution worker 1 may get the first block of random numbers and worker 2 the second block, while in another execution the situation could be reversed. In fact, any permutation of the block distribution

can take place on each iteration of the main loop, and the model checker will be forced to explore all of them, leading to a rapid blowup in the number of states of M_{par}. This problem does *not*, however, arise for M_{seq}, because M_{seq} utilizes the random numbers in a fixed order. Moreover, for all executions of M_{seq}, both the output vector and the path condition turn out to be invariant under these permutations. The upshot of our symmetry reduction theorem is that in these circumstances, it suffices to explore only one of these permutations in M_{par}, rather than all of them.

Using these reductions, we were able to establish IEEE equivalence for $n \leq 9$.

4. RELATED WORK

The idea of representing computations symbolically has a long history and has enjoyed many applications, including to testing and debugging (e.g., [2,3,8]). There has also been some work incorporating these ideas into model checking. For example, a component of the SLAM toolkit [1] translates a C program into a program that operates solely on Boolean variables corresponding to predicates in the original program. A theorem prover is used in that process to determine the effect of each statement in the original program on the predicates. Another component uses symbolic execution to determine whether a path through the Boolean program corresponds to an actual execution of the original program. This is similar in spirit to our method, which translates a program into one which operates on symbolic expressions and uses a (very lightweight) form of theorem proving to determine branches and expression equivalence.

Symbolic execution has also been used to improve the precision of Java PathFinder, in order to verify properties of Java programs that manipulate complex data structures and that may even contain unbounded loops [13, 17].

Our approach differs from this previous work in several ways: (1) in the way we use the path condition to filter out executions of the parallel program that are not consistent with a sequential execution, (2) in our emphasis on complex floating-point expressions, rather than on heap-allocated data and integer expressions, and (3) in our use of the value numbering scheme to represent the state space efficiently.

There are a number of tools and techniques that can be used to estimate the error arising from floating-point computations in programs; see [14] for a description and comparison of some of these.

5. CONCLUSIONS AND FUTURE WORK

We have described a method that uses model checking techniques in combination with symbolic execution to verify the correctness of the calculations performed by parallel programs—even complex floating-point calculations. We have successfully applied this method to four quite different examples, scaling to configurations of between 6 and 17 processes. While these numbers are much smaller than those that arise in practice, evidence from the application of model checking techniques with other kinds of software suggests that problems are usually exposed by verification of relatively small configurations. This is quite different from the case with testing, where the small size may make it difficult to trigger particular pathological patterns of behavior. The difference is due to the fact that model checking takes into account all possible executions of the model.

The key idea of our method is to compare a sequential and a parallel program. This approach takes advantage of the fact that, since it is usually easier to construct a correct sequential numerical program, developers often start with a sequential version or develop one in tandem with the parallel version. The effectiveness of our method relies on the assumption that the computations performed in the two programs are close to being exactly the same, though those computations may be distributed in a complex way in the parallel program. This assumption means that it is usually relatively inexpensive to determine if two symbolic expressions are equivalent or if one symbolic predicate implies another. The further removed the computations in the two programs become, the more powerful the symbolic manipulation must be in order to arrive at a conclusive result. If the computations performed by the two programs are very different, we might argue that the sequential program is not a good specification for the parallel one. Nevertheless, as we examine more complex programs, it is certainly possible that the kind of lightweight symbolic manipulation and theorem proving that we are currently using will no longer suffice. For this reason, we are exploring ways to integrate our approach with more sophisticated symbolic algebra and theorem proving tools.

One of the biggest barriers to the wide adoption of model checking techniques is the problem of model extraction. For our preliminary study, we built Spin models by hand, and though we designed our symbolic package to make this as easy as possible, it still requires a great deal of expertise. The ideal situation would be to have tools that automatically extract the models from source code, and indeed a great deal of research on this subject has been carried out, at least for other domains. We are exploring ways to adapt these techniques to MPI programs, though we expect to encounter some significant challenges when it comes to automatically creating models of programs with large amounts of floating-point data.

Perhaps the greatest problem with model checking is *state explosion*: the fact that the number of states of a program typically grows exponentially with the number of processes. A vast array of techniques has been developed to counteract this problem, and we have demonstrated that some of these, such as partial order and symmetry reductions, can be adapted to work with our method. Spin turned out to be an excellent platform for the rapid prototyping of our method, although it was too difficult to code some of the optimizations that we wanted to consider. We plan to explore these using the Bogor model checker [18], which is designed to allow easy customizations of its search strategy and other components. We intend to try to take advantage of this platform to explore a wide range of optimizations and extensions to our method.

Acknowledgments

This material is based upon work supported by the National Science Foundation under awards CCF-0427071 and CCR-0205575 and by the U.S. Department of Defense/Army Research Office under awards DAAD19-01-1-0564 and DAAD-19-03-1-0133. We also wish to thank the Computing Research Association's CRA-W Distributed Mentor Program and the College of Natural Science and Mathematics at the University of Massachusetts for sponsoring the program and

funding, respectively, for Mironova during the summer of 2004.

Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation or the U.S. Department of Defense/Army Research Office.

6. REFERENCES

[1] T. Ball and S. K. Rajamani. Automatically validating temporal safety properties of interfaces. In M. B. Dwyer, editor, *Model Checking Software: 8th International SPIN Workshop, Toronto, Canada, May 19–20, 2001, Proceedings*, volume 2057 of *Lecture Notes in Computer Science*, pages 103–122. Springer, 2001.

[2] R. S. Boyer, B. Elspas, and K. N. Levitt. SELECT—a formal system for testing and debugging programs by symbolic execution. In *Proceedings of the International Conference on Reliable Software*, pages 234–245. ACM Press, 1975.

[3] L. A. Clarke. A system to generate test data and symbolically execute programs. *IEEE Transactions on Software Engineering*, 2(3):215–222, 1976.

[4] R. Cousot, editor. *Verification, Model Checking, and Abstract Interpretation: 6th International Conference, VMCAI 2005, Paris, January 17–19, 2005, Proceedings*, volume 3385 of *Lecture Notes in Computer Science*, 2005.

[5] D. Goldberg. What every computer scientist should know about floating-point arithmetic. *ACM Computing Surveys*, 23(1):5–48, Mar. 1991.

[6] S. Graf and L. Mounier, editors. *Model Checking Software: 11th International SPIN Workshop, Barcelona, Spain, April 1–3, 2004, Proceedings*, volume 2989 of *Lecture Notes in Computer Science*. Springer, 2004.

[7] W. Gropp, E. Lusk, and A. Skjellum. *Using MPI: Portable Parallel Programming with the Message-Passing Interface*. MIT Press, 1999.

[8] S. L. Hantler and J. C. King. An introduction to proving the correctness of programs. *ACM Computing Surveys*, 8(3):331–353, 1976.

[9] G. J. Holzmann. *The* Spin *Model Checker*. Addison-Wesley, 2004.

[10] IEEE. 754-1985 IEEE standard for binary floating-point arithmetic, 1985.

[11] IEEE. 854-1987 IEEE standard for radix-independent floating-point arithmetic, 1987.

[12] G. E. Karniadakis and R. M. Kirby II. *Parallel Scientific Computing in C++ and MPI*. Cambridge University Press, 2003.

[13] S. Khurshid, C. S. Păsăreanu, and W. Visser. Generalized symbolic execution for model checking and testing. In H. Garavel and J. Hatcliff, editors, *Tools and Algorithms for the Construction and Analysis of Systems, 9th International Conference, TACAS 2003, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2003, Warsaw, Poland, April 7-11, 2003, Proceedings*, volume 2619 of *Lecture Notes in Computer Science*, pages 553–568. Springer, 2003.

[14] M. Martel. An overview of semantics for the validation of numerical programs. In Cousot [4], pages 59–77.

[15] Message Passing Interface Forum. MPI: A Message-Passing Interface standard, version 1.1. `http://www.mpi-forum.org/docs/`, 1995.

[16] Message Passing Interface Forum. MPI-2: Extensions to the Message-Passing Interface. `http://www.mpi-forum.org/docs/`, 1997.

[17] C. S. Păsăreanu and W. Visser. Verification of Java programs using symbolic execution and invariant generation. In Graf and Mounier [6], pages 164–181.

[18] Robby, M. B. Dwyer, and J. Hatcliff. Bogor: an extensible and highly-modular software model checking framework. In *ESEC/FSE-11: Proceedings of the 9th European Software Engineering Conference held jointly with the 11th ACM SIGSOFT International Symposium on Foundations of Software Engineering*, pages 267–276, Helsinki, Finland, 2003. ACM Press.

[19] J. J. Rotman. *An Introduction to the Theory of Groups*. Springer, New York, fourth edition, 1995.

[20] S. F. Siegel. Efficient verification of halting properties for MPI programs with wildcard receives. Technical Report UM-CS-2004-77, Department of Computer Science, University of Massachusetts, 2004.

[21] S. F. Siegel. Efficient verification of halting properties for MPI programs with wildcard receives. In Cousot [4], pages 413–429.

[22] S. F. Siegel and G. S. Avrunin. Verification of MPI-based software for scientific computation. In Graf and Mounier [6], pages 286–303.

[23] S. F. Siegel and G. S. Avrunin. Modeling wildcard-free MPI programs for verification. In *Proceedings of the 2005 ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming: PPoPP'05, June 15–17, 2005, Chicago, Illinois, USA*, pages 95–106. ACM Press, 2005.

APPENDIX

A. REDUCTION THEOREMS

A.1 A partial order reduction theorem for MPI programs

To explore all possible terminal *states* of a program, it is usually not necessary to explore all possible *paths* through the program. That is because there are often many *equivalent* paths, i.e., paths terminating in the same terminal state. This means that we may restrict the set of paths that we explore, as long as we are sure to keep at least one representative from each equivalence class. The theorem below justifies such a restriction for MPI programs.

To state the theorem, we adopt the notation of [21]. Let $\mathcal{M}$ be any model of an MPI program, and ν a *channel size vector*. Thus ν assigns to each channel c a value $\nu(c)$ which is either ∞ or a non-negative integer. Suppose that T is a finite trace of $\mathcal{M}$ from a global state σ_0 to a global state σ_f and that T is *ν-bounded*, i.e.,

$$\forall c \in \mathsf{Chan}\colon\ \mathsf{maxlen}_c(T) \leq \nu(c).$$

Recall that this means that the number of messages queued in any channel c never exceeds the bound $\nu(c)$. Suppose

further that σ_f is *ν-halted*, i.e., there is no global transition departing from σ_f that does not cause the number of queued messages in some c to exceed $\nu(c)$. Another way of saying this is that there is no way to extend T to a ν-bounded trace of length $n+1$, where $n = |T|$.

We consider a game that involves two players, the *scheduler* and the *selector*, and constructs a new ν-bounded trace T'. The game consists of a sequence of stages, each of which extends the trace by one transition. So, at the beginning of stage i, the first $i-1$ transitions of T' have already been selected. Stage i proceeds as follows. First, if T' terminates in a ν-halted state, then the game *ends in deadlock*. Otherwise, the scheduler chooses a non-empty subset of enabled transitions, according to a certain rule that we describe below. The selector then chooses a specific transition from this subset and appends it to T'. The game ends if it deadlocks, or after n stages have completed, whichever occurs first. The selector wins the game if the final trace constructed terminates in σ_f, else the scheduler wins. The theorem states that there is a strategy for the selector so that the selector always wins.

We now describe the *scheduling rule* that constrains the choice made by the scheduler. Say the current state is σ. Let E_σ denote the set of all *ν-enabled* global transitions at σ, i.e., the set of all global transitions enabled at σ that do not cause the number of queued messages for any channel c to exceed its bound $\nu(c)$. Suppose E_σ is non-empty. Then the scheduler must choose a non-empty subset F of E_σ that satisfies at least one of the following conditions: (1) $F = \{\tau\}$, where τ is not a local event transition, nor a receive transition emanating from a wildcard receive state (i.e., a state from which there are outgoing transitions labeled by receives on at least two distinct channels), nor a synchronous transition for which the associated receive transition emanates from a wildcard receive state, (2) F is the set of all transitions enabled in a single process that is at a local event state, or (3) $F = E_\sigma$. We call such a set F an *acceptable set at σ*, and the set of all acceptable sets at σ will be denoted C_σ.

To state the theorem precisely, we introduce the following notation. First, for any finite sequence $S = (x_1, \ldots, x_n)$ of elements of a set X, and any $x \in X$, we let $S \cdot x$ denote the sequence obtained by appending x to the end of S. Next, given any two global states σ and σ_f of $\mathcal{M}$, we define statements $\theta(\sigma, \sigma_f, n)$, for all $n \geq 0$, as follows: $\theta(\sigma, \sigma_f, 0)$ is the statement $\sigma = \sigma_f$, and for $n > 0$, $\theta(\sigma, \sigma_f, n)$ is the statement

$$E_\sigma \neq \emptyset \wedge \forall F \in C_\sigma \; \exists \tau \in F \colon \theta(\mathsf{des}(\tau), \sigma_f, n-1).$$

This is a formal way of stating that, starting from σ, the selector can always force the trace to terminate at σ_f in n steps, no matter what moves are made by the scheduler.

THEOREM 1. *Let $\mathcal{M}$ be a model of an MPI program, ν a channel size vector, T a ν-bounded trace in $\mathcal{M}$ from a global state σ to a ν-halted global state σ_f. Then $\theta(\sigma, \sigma_f, |T|)$ holds.*

PROOF. The proof is by induction on $|T|$. If $|T| = 0$ then $\sigma = \sigma_f$, and since this is exactly the statement $\theta(\sigma, \sigma_f, 0)$, the theorem holds. So suppose $n > 0$, $T = (\tau_1, \ldots, \tau_n)$, and the theorem holds for any trace of length less than n. We cannot have $E_\sigma = \emptyset$ since τ_1 is ν-enabled at σ. Let $F \in C_\sigma$.

If $F = E_\sigma$ then let $\tau = \tau_1$. Then $(\tau_2, \ldots, \tau_n)$ is a ν-bounded trace from $\mathsf{des}(\tau)$ to σ_f of length $n-1$, and so by the inductive hypothesis, $\theta(\mathsf{des}(\tau), \sigma_f, n-1)$ holds. Hence $\theta(\sigma, \sigma_f, |T|)$ holds.

Suppose F is the set of all ν-enabled local event transitions from a single process p. Now there must exist an integer i such that $1 \leq i \leq n$ and τ_i is in process p. For if not, there would still be a local event transition in p enabled at σ_f, and σ_f would not be a ν-halted state. Let i be the least integer with this property. We claim that there is a ν-bounded trace

$$T' = (\tau_i', \tau_1', \ldots, \tau_{i-1}', \tau_{i+1}', \ldots, \tau_n')$$

from σ to σ_f with $\mathsf{label}(\tau_j') = \mathsf{label}(\tau_j)$ for all j. This is because we may move τ_i to the left one step at a time, applying [20, Lemma 1] at each step. Let $\tau = \tau_i'$, and argue as in the paragraph above to see that $\theta(\sigma, \sigma_f, |T|)$ holds.

Suppose F is a singleton set $\{\tau\}$, where τ is either a send, receive, or synchronous transition. Say τ is a receive in process p. Then, according to the scheduling rule, at σ, p must be in a receiving state for a sole channel c, and so $\mathsf{label}(\tau) = c?x$ for some x that is already queued at σ. Now there must exist some i such that (1) $1 \leq i \leq n$, (2) τ_i is a receive in process p, and (3) there is at most one j such that $1 \leq j < i$ and τ_j belongs to process p, and, if there is such a j, then σ is a send-receive state and τ_j is the send emanating from that state. The reason for this is that if it were not the case, there would still be a receive enabled at σ_f, and σ_f would not be ν-halted. Now we argue as before to move τ_i to the left, using [20, Lemma 2] to move past the send τ_j if necessary. The only thing we must check is that the message received by τ_i was already queued at σ. However, since c is the sole receiving channel for p at σ, τ_i must also be a receive on c, and so we must have $\mathsf{label}(\tau_i) = c?x$, as required. Now we proceed to argue as in the paragraph above that $\theta(\sigma, \sigma_f, |T|)$ holds. The case where τ is a send is similar but easier since we do not have to deal with wildcards. If τ is a synchronous transition, then first decompose it into its send and receive parts, then move each all the way to the left, and then recompose them into a synchronous transition. □

We now examine some practical consequences of Theorem 1. Say we are using SPIN to verify an assertion on the terminal states of $\mathcal{M}$. In creating a SPIN model, we must specify a finite bound $\nu(c)$, for each channel c, when c is declared. By the *scheduling policy* we mean the mechanism of SPIN that determines the set of all possible next transitions from a given state. The scheduling policy plays the role of the scheduler in our game. In its default mode, the scheduling policy returns all ν-enabled transitions. Hence in the default mode, SPIN will explore all ν-halted states that are reachable by ν-bounded traces from the initial state.

Notice that, if there are terminal states that can only be reached by traces in which the number of queued messages for some c exceeds $\nu(c)$, these states will not be explored by SPIN. In some cases, one may verify that there are no such states by using assertions to check that control never reaches a send statement for a channel when that channel is full. In other cases, this may not be possible, or the channel sizes required may be so large that the verification becomes infeasible. In such cases we may still use a less-than-satisfactory channel size and satisfy ourselves with a result that is not quite conservative. (Of course, if SPIN finds an error, this is just as helpful as if we had used unbounded channels.) The situation is similar to the need that sometimes arises to place small bounds on the number of loop iterations, and

is a problem that often arises with finite-state verification techniques. What we will see shortly, however, is that the reduction strategy we describe cannot make the problem any worse, i.e., if there exists a property violation within the specified bounds, it will still be found after applying the reduction.

Now, the scheduling policy used by SPIN can be restricted with the careful use of `atomic` blocks. When control is inside an `atomic` block of a process, SPIN's scheduling policy returns only the ν-enabled transitions from that process, assuming there is at least one. If there are none, then the process loses atomicity, and the scheduling policy returns the set of all ν-enabled transitions. A special rule is used for *rendezvous channels*, i.e., channels of size 0. A send on a rendezvous channel is not blocked precisely when the receiving process is in a position to receive the message; in this case, the SPIN scheduling policy returns just the synchronous transition and control passes to the receiving process. If the receiving process happens to also be inside an `atomic` block, then atomic control passes directly from the sender to the receiver.

Now we can use `atomic` statements in any way, as long as the resulting scheduling policy obeys the scheduling rule defined above. Let us say, for example, that we have inserted `atomic` statements in such a way that if a wildcard receive occurs within an `atomic` block B then it must be the first statement in B and B cannot be inside another `atomic` block. Assume also that $\nu(c) \geq 1$ for all c, so there are no rendezvous channels. These assumptions mean that the only state from which SPIN's scheduling policy can select a wildcard receive is one in which no process has atomic control. But if no process has atomic control, SPIN's scheduling policy will return the set of all ν-enabled transitions. Hence the scheduling policy satisfies the scheduling rule, and we are guaranteed that SPIN will still explore all ν-halted states reachable from the initial state by ν-bounded traces.

Consider now the case where $\mathcal{M}$ has no wildcard receives, and let ν be any channel size vector. Say that we construct a SPIN model of $\mathcal{M}$ in which we set all channel sizes to 0 and place the code of each process in a single `atomic` block. What is the resulting scheduling policy? In any state, it will return either (1) a singleton set consisting of a synchronous transition (that by assumption does not involve a wildcard), or (2) the set of all local transitions in a single process, or (3) the empty set, if the state is potentially halted (i.e., no synchronous or local event transition is enabled). If (3) occurs when the state is not terminal, SPIN will report this as an improper end state (i.e., a deadlock). Hence if the search returns without ever reporting an improper end state, then the scheduling policy satisfies the scheduling rule, and we are guaranteed that the search has visited every ν-halted state of $\mathcal{M}$ reachable by a ν-bounded trace.

A.2 A symmetry reduction theorem for MPI programs

In this section we prove a theorem that has consequences for numerical programs in which the output vectors and path conditions exhibit symmetry in the symbolic constants. The theorem is expressed using the language of group theory and group actions [19].

Let n and m be non-negative integers, G a finite group, and X a G-set of cardinality n. Let

$$\mathcal{X} = \{(x_1, \ldots, x_n) \mid \{x_1, \ldots, x_n\} = X\},$$

which is a set of cardinality $n!$. Let Π and Y be G-sets, and let $\mathcal{Y} = Y^m$. The action of G on X induces an action of G on $\mathcal{X}$ by defining

$$g(x_1, \ldots, x_n) = (gx_1, \ldots, gx_n),$$

for $g \in G$. The action of G on Y induces a component-wise action on $\mathcal{Y}$ as well. We call the 6-tuple

$$\mathcal{C} = (G, X, \mathcal{X}, \Pi, Y, \mathcal{Y})$$

a *symbolic context*.

Let $\mathcal{C}$ be a symbolic context. Consider a pair of functions (C, f), where C assigns, to each $\mathbf{x} \in \mathcal{X}$ and $p \in \Pi$, a set $C(\mathbf{x}, p)$, and f assigns to each triple $(\mathbf{x}, p, c)$, where $\mathbf{x} \in \mathcal{X}$, $p \in \Pi$, and $c \in C(\mathbf{x}, p)$, an element $\mathbf{y} = f(\mathbf{x}, p, c) \in \mathcal{Y}$. Assume that for all $g \in G$, $\mathbf{x} \in \mathcal{X}$, $p \in \Pi$, and $c \in C(\mathbf{x}, p)$, the following both hold:

$$C(g\mathbf{x}, gp) = C(\mathbf{x}, p) \tag{1}$$

$$gf(\mathbf{x}, p, c) = f(g\mathbf{x}, gp, c). \tag{2}$$

Then we call (C, f) a *symbolic model* over $\mathcal{C}$.

THEOREM 2. *Let $\mathcal{C} = (G, X, \mathcal{X}, \Pi, Y, \mathcal{Y})$ be a symbolic context and let (C, f) be a symbolic model over $\mathcal{C}$. Suppose there are $\mathbf{x} \in \mathcal{X}$, $p \in \Pi$, and $\mathbf{y} \in \mathcal{Y}$ for which the following all hold:*

1. *$gp = p$ for all $g \in G$,*
2. *$g\mathbf{y} = \mathbf{y}$ for all $g \in G$, and*
3. *$\forall c \in C(\mathbf{x}, p)\ \exists g \in G\colon f(g\mathbf{x}, p, c) = \mathbf{y}$.*

Then $f(\mathbf{x}, p, c) = \mathbf{y}$ for all $c \in C(\mathbf{x}, p)$.

PROOF. Given c, choose g to satisfy hypothesis 3. Then

$$\begin{aligned} f(\mathbf{x}, p, c) &= g^{-1} f(g\mathbf{x}, gp, c) \\ &= g^{-1} f(g\mathbf{x}, p, c) \\ &= g^{-1} y \\ &= \mathbf{y}. \quad \square \end{aligned}$$

Now we describe the application of this theorem to symbolic models of numerical programs. In this application, the set X is the set of symbolic constants, and $\mathcal{X}$ is the set of all possible input vectors to the model. The group G may be any subgroup of Σ_X, the group of all permutations of X. The set Π is the set of all boolean-valued symbolic expressions in the symbolic constants, e.g., $(x_1 x_2)/x_3 \geq 0 \wedge x_2 \neq x_3$. A path condition, for example, is an element of Π. Notice that the action of G on X extends naturally to an action on Π in which G acts trivially on operators and literals. We may also take Π to be the set of all boolean-valued symbolic expressions modulo an operation-preserving equivalence relation $\sim$, as long as $\sim$ is preserved by the action of G, i.e., $p \sim q \Rightarrow gp \sim gq$ for all $g \in G$. The examples of equivalence relations that we have considered in this paper all satisfy this property. The set Y is the set of all real-valued symbolic expressions in the symbolic constants, e.g., $(x_1 x_2)/x_3 + x_4$. Again, we may apply an appropriate equivalence relation. The set $\mathcal{Y}$ is the set of all symbolic output vectors.

The pair (C, f) represents our numerical program P. The set $C(\mathbf{x}, p)$ corresponds to the set of all executions of P on input $\mathbf{x}$ that are consistent with the path condition p. The element $\mathbf{y} = f(\mathbf{x}, p, c)$ represents the output of P when given input $\mathbf{x}$, a path condition p, and a particular execution c

consistent with p. The assumptions (1) and (2) express what is essentially a functorial property in the symbolic constants: permuting the names of the input does not change the set of behaviors of the program, nor does it change the output produced by the program, except to permute the names of the symbolic constants in the output in the same way that they were permuted in the input.

The theorem may now be interpreted as follows: suppose we are given an input vector and a path condition p and output vector $\mathbf{y}$ such that both p and $\mathbf{y}$ are invariant under the action of G. Then for each possible path through the model that is consistent with p, we may first permute the input according to any element of G before computing the output produced by that path. If the output is always $\mathbf{y}$ then we may conclude the output would have been $\mathbf{y}$ even if we had not permuted the input.

B. CODE

B.1 Parallel Gaussian Elimination Code

```
double matrix[M];
                  .
                  .
                  .
int top,col,row,j,rank,nprocs;
double pivot,tmp;
double toprow[M];
MPI_Status status;
MPI_Comm_size(MPI_COMM_WORLD, &nprocs);
MPI_Comm_rank(MPI_COMM_WORLD, &rank);
for (top=col=0; top<N && col<M; top++, col++) {
  for (; col < M; col++) {
    if (matrix[col]!=0.0 && rank>=top)
      MPI_Allreduce(&rank, &row, 1,
        MPI_INT, MPI_MIN, MPI_COMM_WORLD);
    else
      MPI_Allreduce(&nprocs, &row, 1,
        MPI_INT, MPI_MIN, MPI_COMM_WORLD);
    if (row<nprocs) break;
  }
  if (col>=M) break;
  if (row!=top) {
    if (rank==top)
      MPI_Sendrecv_replace(matrix, M, MPI_DOUBLE,
        row, 0, row, 0, MPI_COMM_WORLD, &status);
    else if (rank==row)
      MPI_Sendrecv_replace(matrix, M, MPI_DOUBLE,
        top, 0, top, 0, MPI_COMM_WORLD, &status);
  }
  if (rank==top) {
    pivot = matrix[col];
    for (j=col; j<M; j++) {
      matrix[j] /= pivot;
      toprow[j] = matrix[j];
    }
  }
  MPI_Bcast(toprow, M, MPI_DOUBLE, top,
    MPI_COMM_WORLD);
  if (rank!=top) {
    tmp = matrix[col];
    for (j=col; j<M; j++)
      matrix[j] -= toprow[j]*tmp;
  }
}
```