ABSTRACT
We introduce a methodology, based on symbolic execution, for Concurrent Bounded Model Checking. In our approach, we translate a program into a formula in a disjunctive form. This design enables concurrent verification, with a main thread running symbolic execution, without any constraint solving, to build subformulas, and a set of worker threads running a decision procedure for satisfiability checks.

We have implemented this methodology in a tool called JCBMC, the first bounded model checker for Java. JCBMC is built as an extension of Java Pathfinder, an open-source verification platform developed by NASA. JCBMC uses Symbolic PathFinder (SPF) for the symbolic execution, Z3 as the solver and implements concurrency with multi-threading.

For evaluation, we compare JCBMC against SPF and the Bounded Model Checker CBMC. The results of the experiments show that we can achieve significant advantages of performance over these two tools.

Categories and Subject Descriptors
D.1.3 [Concurrent Programming]: Parallel programming; D.2.4 [Software/Program Verification]: Formal methods

General Terms
Algorithms, Performance, Theory, Verification

Keywords
Bounded Model Checking; Concurrency; Symbolic Execution

1. INTRODUCTION
Bounded Model Checking (BMC) is a popular verification technique that works by unwinding the transition relation of a program for a fixed number of steps $k$ and checking whether a property violation can occur in $k$ or fewer steps. This property checking is then performed by a SAT or SMT solver. BMC is widely used in the hardware industry.

For software, the application of BMC for ANSI-C is embodied in the CBMC tool. In CBMC, a C program containing assertions is encoded into a formula (in Static Single Assignment form), which is then fed to a SAT solver or an SMT solver to check its satisfiability. A satisfying assignment indicates that an error was found. The formula is in conjunctive normal form.

BMC has not been explored so far for many other languages, including Java. However, explicit-state model checking tools such as Java Pathfinder (JPF) have been very effective for the verification of many Java applications. Furthermore, there has been an explosion of symbolic execution tools that have been used successfully for many languages. In particular, relevant to our approach, Symbolic PathFinder (SPF) is a symbolic execution tool built as an extension of JPF, that provides a symbolic analysis for Java programs.

In this paper, based on our previous work, we propose a new methodology for BMC which is based on “classical” symbolic execution (SE) in the sense of King. The way CBMC transforms a program into Static Single Assignment form can also be viewed as executing the program symbolically. However, this encoding is different from the Symbolic Execution of King and we evaluate the two encodings as part of the work reported here. Our methodology is not language specific and only relies on a symbolic executor for that language and an SMT solver.

By using SE we obtain a translation of a program and assertions into a disjunctive formula encoding the path conditions for each bounded path explored in the code. This suggests a simple concurrent verification strategy that relies on the observations that any subset of disjuncts can be separately checked for satisfiability and whenever a subset is found to be satisfiable the satisfiability task can be stopped. Hence the verification of disjunctive formulas is naturally parallelizable.

We have implemented this methodology in a tool JCBMC, a Java Concurrent Bounded Model Checker, which uses SPF to generate the disjunctive formula from the code (constraint solving is turned off in SPF itself) and while generating the formula it sends sub-formulas to multiple worker threads for satisfiability checking. JCBMC handles programs with multi-threading and recursive input data structures and relies on a standard SMT solver, namely Z3, for solving the constraints. Other solvers can be incorporated and different solvers can be used for solving different path constraints in parallel. Although JCBMC is only a prototype, its performance, compared with existing tools, i.e. SPF and CBMC, is remarkable. We summarize our contributions as follows:

- A methodology for concurrent bounded model checking that is based on “classical” SE and it is naturally parallelizable.
- The methodology is language independent and supports assume-guarantee reasoning.
- A tool JCBMC, a concurrent bounded model checker for Java.
- Experiments to show effectiveness of the tool for verification of programs with multi-threading and data structures.
- Comparisons with bounded model checking and “classical” SE, as embodied by CBMC and SPF respectively.
2. BACKGROUND

A program $P$ is modelled as a transition system:

$$P = (S, I, F, T)$$

where $S$ is the set of program states; $I \subseteq S$ is the set of initial states; $F \subseteq S$ is the set of final states; and $T \subseteq S \times S$ is the transition relation. Under this setting, a trace of (a concrete) execution of the program $P$ is represented by a sequence of states:

$$\rho = s_0s_1..s_k$$

such that $s_0 \in I, s_k \in F$ and $\langle s_i, s_{i+1} \rangle \in T$ for all $i \in \{0, ..., k-1\}$.

2.1 Bounded Model Checking and CBMC

A trace can be also seen in logical form: the set $I$ and the relation $T$ can be written as their characteristic functions: $s_0 \in I$ iff $I(s_0)$ holds; $\langle s_i, s_{i+1} \rangle \in T$ iff $T(s_i, s_{i+1})$ holds. In this way, a trace $\rho$ is represented by the formula:

$$I(s_0) \land \bigwedge_{i=0}^{k-1} T(s_i, s_{i+1})$$

Clearly the transition system $P$ is a model for such a formula, i.e. $P$ is a model for all formulas representing traces of the program. The aim of BMC is to find bugs or prove their absence up to some bounded $k$ number of transitions. That means it explores all traces $\rho = s_0s_1..s_k$ of the program $P$, in which $s_k$ needs not to be in $F$. Notice that because of the bound $k$ there are only a finite number of traces to explore hence we can represent the bounded program as a formula $\mathcal{C}$ which is a conjunction of formulas, whose conjoints are possible traces. Notice formulas can also represent symbolic traces, for example if in a formula the value of a program variable is left unspecificed then there can be several concrete traces satisfying that formula. Formulas satisfied by set of concrete traces can be referred to as symbolic traces.

CBMC translates a C program into a logical formula $\mathcal{C}$ which is then used as a model for the property $P$ to be verified. The property is verified by the C program iff $\mathcal{C} \land P$ is valid. This can be checked by a satisfiability solver on $\mathcal{C} \land \neg P$. In fact if $\mathcal{C} \land \neg P$ is true in the model then one trace will satisfy $\neg P$ hence the property is not valid. On the other hand if $\mathcal{C} \land \neg P$ is false in the model then no trace will satisfy $\neg P$ hence $P$ is valid.

2.2 Symbolic PathFinder: Symbolic Execution for Java Bytecode

Symbolic PathFinder (SPF) is a symbolic execution framework built on top of the Java PathFinder (JPF) model checking tool-set for Java bytecode analysis. SPF implements a Bytecode interpreter that replaces the standard, concrete execution semantics of bytecodes with a non-standard symbolic execution. Non-deterministic choices in branching conditions are handled by means of JPF’s choice generators. Each non-deterministic choice has associated a path condition. JPF’s listeners are used to monitor and influence the symbolic execution and to collect and print its results. Symbolic execution of looping programs may result in an infinite symbolic execution tree; for this reason, SPF is run with a user-specified bound on the search depth. By default, whenever the path condition is updated, SPF invokes an off-the-shelf solver to check its satisfiability; if the path condition is found to be unsatisfiable, SPF backtracks. As a result, by default, SPF explores only feasible paths. Note that SPF also has an option to run with no solving (that we use in our work here); as a result of this option, SPF will explore all the possible paths (feasible and infeasible) through the program, up to the given bound. SPF uses lazy initialization [29] to handle dynamic input data structures (e.g., lists and trees). Multi-threading is handled systematically using the search mechanisms in JPF core.

3. CONCURRENT BOUNDED MODEL CHECKING

Our method for concurrent bounded model checking is illustrated in Fig. 1. The inputs are: a program under test, a property to verify and three parameters – $B$ is the search bound, $N$ is the number of workers and $D$ is the number of disjuncts to give each worker. The goal is to check if the property holds in the program, up to bound $B$.

![Figure 1: Concurrent BMC architecture](image)

The program under test is analysed using the “classical” (bounded) SE procedure with constraint solving turned off. This means whenever a path condition is updated, we do not check its satisfiability, but rather continue the exploration. As a result, SE may explore infeasible paths, which will be checked later using constraint solving. Our approach can be used for the bounded verification of safety properties, which we assume have been reduced to checking assertions embedded in the code. Furthermore, our method supports both assume and assert statements to enable assume-guarantee style verification. The assumed conditions are simply added to the path conditions during the symbolic execution.

The result of SE is a disjunction of path conditions, encoding constraints on the inputs to follow those paths, up to the pre-specified search bound. Among these paths, only the ones that may lead to assert violations are selected for solving. This is achieved by the Controller which collects sets of $D$ violating path conditions and sends them for solving to parallel worker threads, using off-the-shelf solvers. The workers start solving as soon as they receive the disjunctive formulas, which may happen while the symbolic execution is still exploring the program. The verification terminates as soon as one of the threads finds a satisfying assignment, in which case an error is reported, or when all the disjunctions are found to be un-satisfiable, in which case the assertion holds (no error) up to the given bound. Note that if the symbolic executor discovers no potentially violating paths (i.e. the error is unreachable), then no solving will be performed.

In general, we use SE to explore all possible symbolic paths up to a certain length, and then encode the program together with the property to check into a formula of the form: $\bigvee_{i=0}^{D-1} pc_i$ where $M$ is the number of paths that may trigger the error. This form allows us to divide the formula into blocks of $D$ disjunctions:

$$\bigvee_{i=0}^{D-1} pc_i \lor \bigvee_{i=D}^{2D-1} pc_i \lor \cdots \lor \bigvee_{i=(k-1)D}^{kD-1} pc_i \lor \bigvee_{i=kD}^{M} pc_i$$
In this way, we can solve the formula concurrently using several threads, each one solving a single block. A model of a single block is also a model of the formula, therefore the procedure stops when any of the threads find out a model. In JCBMC, after the main thread generates a sub-formula and passes it to a worker thread, it moves on to generate the next sub-formula, while the worker thread solves the given sub-formula concurrently.

3.1 Comparing our approach with Bounded Model Checking and Symbolic Execution

Compared with classical BMC we use an explicit enumeration of paths, while BMC uses an implicit enumeration of paths. Although at first glance the implicit encoding should be better our experiments, even with sequential JCBMC, show that this is not the case. Furthermore, the explicit enumeration is easily parallelizable, with simple and natural load balancing for different threads. Crucially our approach stops as soon as a path leading to an error is found to be satisfiable, while with classical BMC, all the program needs to be explored.

Compared with classical symbolic execution: we solve only in the end. So obviously the price to be paid is the exploration of infeasible paths. On the other hand, again, it is naturally parallelizable and constraint solving, which is one of the bottlenecks in SE, can be done in parallel, even with different solvers, with little coordination, if any, needed.

SPF is used to extract all possible symbolic paths up to a given conditional branching depth. The proposed method outperforms SPF (with solving), even in the sequential version (probably even with one worker). One reason could be that most programs do not contain unreachable symbolic paths, therefore "lazy" checking of complete symbolic paths is more efficient than checking prefixes after each branching. This hypothesis could be validated by applying SPF such a lazy way or using incremental SMT solving.

The proposed method does not check each single symbolic path separately but it checks sets of paths. This way it makes use of the clause learning mechanism of SMT solving, in contrast to SPF. It would be also interesting to consider the exchange of certain learnt clauses between the workers to further speed up the satisfiability checks.

The proposed method outperforms classical BMC because it uses substitution instead of introducing sets of variables for each (relevant) program location. It would be interesting to see whether a classical BMC formulation checked by an SMT solver whose theory solver applies substitution as pre-processing would yield similar results.

4. EVALUATION

Our evaluation comprises cases studies to compare JCBMC with SPF (v6, default configuration) and case studies to compare JCBMC with CBMC. To compare with CBMC we have considered C code whose Java translation is almost literal. By JSBMC we denote the sequential implementation of JCBMC where a single worker is used. Experiments are run on a machine equipped with dual Xeon(R) E5-2670 CPUs. The results are shown in Tables 2 and 3.

1To compare both tools with the same solver in the experiments CBMC will be called with option -smt2, and we will use Z3 for satisfiability checks. Note also that CBMC has an option -z3 to use Z3 in SMT1 format. However, in our case studies, using CBMC with this option is much slower than using with SMT2 as in our experiments.

4.1 Comparing with CBMC and SPF

We evaluate our tool against CBMC and SPF in two classical programs that can be written in both C and Java.

4.1.1 Bubble Sort

We consider the classical bubble sort algorithm, which has already been studied in the BMC community. Here, differently from CBMC, we consider the more challenging symbolic version where the values of the array are non-deterministically chosen. We consider both the verification of the assertion “the elements of the array are ordered after bubble sort” and its negation “the elements of the array are not ordered after bubble sort”. We analyse a program implementing bubble sort. It will hence contain no bugs for the positive assertion and will be buggy for the negation. Results are shown in Fig. 2. We notice that while CBMC is better for the positive assertion, JCBMC outperforms the other tools for the negative assertion and is capable of find a counterexample for array sizes of a higher order of magnitude.

4.1.2 Sum of array

We consider the array case studies from the Software Verification competition 2014. In particular sum_array_safe.c for verification and sum_arrayUnsafe.c for refutation. The array size is set to 1000. Results show both SPF and JCBMC outperform CBMC for the unsafe version, while CBMC has a slight advantage for the safe version.

4.2 Comparing with SPF (Java code)

The following examples consist of substantial Java code which is not naturally translatable in C; we hence compare JCBMC only with SPF. Notice JCBMC and SPF are both extensions of JPF: in the case all inputs are concrete they both reduce to JPF-core hence their performance is identical. Hence we only consider programs with symbolic inputs.

4.2.1 Flap controller

This case study is shipped with the distribution of SPF. It is a multi-threaded program modelling a simplified flap controller on an aircraft. It contains 3 classes, and 80 lines of code.

4.2.2 Red Black Tree

This is another example from the SPF distributions (3474 LOC in one class). We check for consistency of the tree after performing put, remove, get and firstKey symbolically.

4.2.3 MER Arbiter

The MER Arbiter models a component of the flight software for NASA JPL’s Mars Exploration Rovers (MER). The MER Arbiter has been modelled in Simulink/Stateflow and it was automatically translated into Java using the Polyglot framework and analyzed with SPF. The configuration for our analysis involved two users and five resources. The example has 268 classes, 553 methods, 4697 lines of code (including the Java Polyglot execution framework) but only approx. 50 classes are relevant. We analyse the code with and without the error (see 6).

2A memory hit is a “run out of memory” problem. This can be addressed by a different memory manager in JPF.

https://github.com/qsphan/jpf-bmc

1To compare both tools with the same solver in the experiments CBMC will be called with option -smt2, and we will use Z3 for satisfiability checks. Note also that CBMC has an option -z3 to use Z3 in SMT1 format. However, in our case studies, using CBMC with this option is much slower than using with SMT2 as in our experiments.
Table:

<table>
<thead>
<tr>
<th>Array size</th>
<th>SPF</th>
<th>JSBMC</th>
<th>CBMC</th>
<th>JCBMC (D = 10)</th>
<th>JCBMC (D = 200)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>5.622</td>
<td>6.204</td>
<td>0.817</td>
<td>1.228</td>
<td>1.389</td>
</tr>
<tr>
<td>30</td>
<td>4m32.790</td>
<td>x</td>
<td>timed out</td>
<td>1.387</td>
<td>2.905</td>
</tr>
<tr>
<td>100</td>
<td>timed out</td>
<td>x</td>
<td>timed out</td>
<td>4.944</td>
<td>34.697</td>
</tr>
</tbody>
</table>

Verification of bubble sort:

<table>
<thead>
<tr>
<th>Array size</th>
<th>SPF</th>
<th>JSBMC</th>
<th>CBMC</th>
<th>JCBMC (D = 10)</th>
<th>JCBMC (D = 200)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>6m19.222</td>
<td>3.712</td>
<td>7.171</td>
<td>4.193</td>
<td>4.262</td>
</tr>
<tr>
<td>6</td>
<td>timed out</td>
<td>26.293</td>
<td>37.816</td>
<td>29.512</td>
<td>21.834</td>
</tr>
<tr>
<td>7</td>
<td>x</td>
<td>x</td>
<td>5m22.641</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>8</td>
<td>x</td>
<td>x</td>
<td>timed out</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Sum of array:

<table>
<thead>
<tr>
<th>Tool</th>
<th>SPF</th>
<th>JSBMC</th>
<th>JCBMC (D = 10)</th>
<th>JCBMC (D = 200)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flap controller (unsafe)</td>
<td>1.043</td>
<td>12.671</td>
<td>1m5.738</td>
<td>1.576</td>
</tr>
<tr>
<td>Red-black tree (safe)</td>
<td>5.602</td>
<td>3.942</td>
<td>3.267</td>
<td>2.774</td>
</tr>
<tr>
<td>MER Arbiter (unsafe)</td>
<td>5.275</td>
<td>8.111</td>
<td>7.479</td>
<td>7.579</td>
</tr>
<tr>
<td>MER Arbiter (safe)</td>
<td>47.063</td>
<td>59.145</td>
<td>57.740</td>
<td>58.886</td>
</tr>
</tbody>
</table>

4.2.4 Discussion

Compared to CBMC, JCBMC scores better in finding counterexamples than in verifying their absence; this is consistent with its design because a counterexample corresponds to a worker thread finding a model of the formula. Compared with SPF, JCBMC can be much better (see bubble sort or red black tree) but can also be comparable or slightly worse (see MER Arbiter and Flap Controller results). The reason for the latter is that the cost of generating path conditions dominates the cost of solving them. Similarly, SPF failed to generate formulas for bubble sort for sizes 7 and higher. Furthermore, an error path (e.g., in MER) may occur at the beginning of the SE exploration, and it is therefore discovered quickly by SPF, while JCBMC still needs to generate the pre-specified number of error paths before solving them. The results suggest one direction for future work, namely to investigate improving the cost of SE-based path generation.

5. RELATED WORK

Related approaches on parallelising BMC [2, 34] address parallel solving of the conjunctive formula that is built for BMC and aim at performing solving at different bounds, where some clauses are shared to enable more efficient SAT solving. In contrast we aim to solve the formulas generated with SE for the same bound, which are naturally disjoint resulting in a simpler parallel algorithm. Furthermore our work aims at verifying programs written in high-level languages such as Java and it is not clear how the previous work, performed in the context of finite state automata, would be applicable. Also related is the work on parallel SAT and SMT solving [21, 22, 35], which can be seen complementing our work; we can use, for example, the parallel version of Z3 [35] in each of the workers to further speed up our proposed approach.

PKIND [19] is a parallel model checker for Lustre that uses k-induction. PKIND runs in parallel the different tasks involved in performing the induction: the base step, the induction step and also the generation of auxiliary invariants used for verification. Thus it performs the parallel work at a higher level of granularity than JCBMC. It would be interesting to investigate if we can replace the parallel tasks in PKIND with our own version of SE-based bounded verification, which in turn is parallelized at the level of granularity of symbolic paths. Parallel model checking has been investigated in the context of explicit-state [2, 22, 24, 17] and symbolic [24] exploration. The latter were done in the context of Binary Decision Diagrams, and hence are very different from ours. These approaches concentrate on partitioning the state space to be explored in parallel and on dealing with the communication overhead between parallel workers. In contrast, in our approach the workers perform the solving independently, with no communication between them.

Previous work on parallel symbolic execution include [32, 21, 11]. All these approaches were done in the context of “classical” or dynamic symbolic execution, using constraint solving during path generation. In contrast the approach we advocate here has a clear separation between path generation and constraint solving, allowing us to easily achieve load balancing between workers, with little communication overhead.

6. CONCLUSION AND FUTURE WORK

We have presented a language independent methodology for concurrent BMC. Based on this methodology we have implemented a concurrent bounded model checker for Java. Future work includes to upgrade its concurrency from single CPU multi-threading to true parallelism and to perform obvious optimisations like replacing SPF with a lighter weight tool, or a parallel version, to reduce the cost of generating path conditions. It will also be interesting to implement the methodology for other languages (C, Python) and investigate how to use SE for IC3 style verification.

7. ACKNOWLEDGEMENTS

We thank Michael Tautschnig for his valuable comments on an earlier version of this paper.

8. REFERENCES


Figure 2: Performance of all tools. “failed” refers to SPF failing to solve the constraints using the integrated solver.

Figure 3: Performance on Flap controller, Red-black tree and MER Arbiter.


Pradeep K. Nalla, J. Weiss, JÅ¡irgrren Ruf, Thomas Kropf, and Wolfgang Rosenstiel. Parallel bounded property checking with symc.


[17] Shahid Jabbar and Stefan Edelkamp. Parallel external