Btor2, BtorMC and Boolector 3.0

Aina Niemetz\textsuperscript{1,2(\Letter)}, Mathias Preiner\textsuperscript{1,2(\Letter)}, Clifford Wolf\textsuperscript{3}, and Armin Biere\textsuperscript{1(\Letter)}

\textsuperscript{1} Johannes Kepler University Linz, Linz, Austria
\textsuperscript{2} Stanford University, Stanford, USA
\textsuperscript{3} Symbiotic EDA, Vienna, Austria

niemetz\texttt{@}cs.stanford.edu

\textbf{Abstract.} We describe Btor2, a word-level model checking format for capturing models of hardware and potentially software in a bit-precise manner. This simple, line-based and easy to parse format can be seen as a sorted extension of the word-level format Btor. It uses design principles from the bit-level format Aiger and follows semantics of the Smt-Lib logics of bit-vectors with arrays. This intermediate format can be used in various verification flows and is perfectly suited to establish a word-level model checking competition. It is supported by our new open source model checker BtorMC, which is built on top of version 3.0 of our SMT solver Boolector. We further provide new word-level benchmarks on which these open source tools are evaluated.

Our format Btor2 generalizes and extends the Btor\textsuperscript{[5]} format, which can be seen as a word-level generalization of the initial version of the bit-level format Aiger\textsuperscript{[2]}. Btor is a format for quantifier-free formulas over bit-vectors and arrays with Smt-Lib\textsuperscript{[1]} semantics but also provides sequential extensions for specifying word-level model checking problems with registers and memories. In contrast to Btor, which is tailored towards bit-vectors and one-dimensional bit-vector arrays, Btor2 has explicit sort declarations. It further allows to explicitly initialize registers and memories (instead of implicit initialization in Btor) and extends the set of sequential features with witnesses, invariant and fairness constraints, and liveness properties. All of these are word-level variants lifted from corresponding features in the latest Aiger format\textsuperscript{[4]}, the input format of the hardware model checking competition (HWMCC)\textsuperscript{[3,6]} since 2011. We provide an open source Btor2 tool suite, which includes a generic parser, random simulator and witness checker. We further implemented a reference bounded model checker BtorMC on top of our SMT solver Boolector. We consider Btor2 as an ideal candidate to establish a word-level hardware model checking competition.

1 Format Description

The syntax of Btor2 is shown in Fig. 1. The \texttt{sort} keyword is used to define arbitrary bit-vector and array sorts. This not only allows to specify multi-dimensional
arrays but can be extended to support (uninterpreted) functions, floating points and other sorts. As a consequence, Btor2 is not backwards compatible with BTOR. For clarity, in Fig. 1 we distinguish between node (line) identifiers ⟨nid⟩ and sort identifiers ⟨sid⟩, and do not allow an identifier to occur in both sets.

Introducing sorts renders type specific keywords such as var, array and acond from BTOR obsolete. Instead, Btor2 uses the keyword input to declare bit-vector and array variables of a given sort. Bit-vector constants are created as in BTOR with the keywords const[dh], one, ones and zero.

Bit-vector and array operators as supported by Btor2 and their respective sorts are shown in Table 1. We use $B^n$ for a bit-vector sort of width $n$, and $\mathcal{I}$ and $\mathcal{E}$ for the index and element sorts of an array sort $\mathcal{A}^{\mathcal{I} \rightarrow \mathcal{E}}$. Note that some bit-vector operators can be interpreted as signed or unsigned. In signed context, as in SMT-LIB, bit-vectors are represented in two’s complement.

### 2 Sequential Extension

As shown in Fig. 1, the sequential extension of Btor2 introduces a state keyword, which allows to specify registers and memories. In contrast to BTOR, where registers are implicitly zero-initialized and memories are uninitialized, Btor2 provides a keyword init to explicitly define initialization functions for states. This enables us to also model partial initialization. For example, initializing a memory with a bit-vector constant zero, zero-initializes the whole memory, whereas
Table 1. Operators supported by BTOR2, where $B^n$ represents a bit-vector sort of size $n$ and $A^{I\rightarrow E}$ represents an array sort with index sort $I$ and element sort $E$.

<table>
<thead>
<tr>
<th>indexed</th>
<th></th>
<th>$B^n \rightarrow B^{n+w}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>[su]ext $w$</td>
<td>(un)signed extension</td>
<td></td>
</tr>
<tr>
<td>slice $u$ $l$</td>
<td>extraction, $n &gt; u \geq l$</td>
<td>$B^n \rightarrow B^{u-l+1}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>unary</th>
<th></th>
<th>$B^n \rightarrow B^n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>not</td>
<td>bit-wise</td>
<td></td>
</tr>
<tr>
<td>inc, dec, neg</td>
<td>arithmetic</td>
<td></td>
</tr>
<tr>
<td>redand, redor, redxor</td>
<td>reduction</td>
<td>$B^n \rightarrow B^n$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>binary</th>
<th></th>
<th>$B^1 \times B^1 \rightarrow B^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>iff, implies</td>
<td>Boolean</td>
<td></td>
</tr>
<tr>
<td>eq, neq</td>
<td>(dis)equality</td>
<td>$S \times S \rightarrow B^1$</td>
</tr>
<tr>
<td>[su]gt, [su]gte, [su]lt, [su]lte</td>
<td>(un)signed inequality</td>
<td>$B^n \times B^n \rightarrow B^1$</td>
</tr>
<tr>
<td>and, nand, nor, or, xnor, xor</td>
<td>bit-wise</td>
<td>$B^n \times B^n \rightarrow B^n$</td>
</tr>
<tr>
<td>rol, ror, sl, sra, srl</td>
<td>rotate, shift</td>
<td>$B^n \times B^n \rightarrow B^n$</td>
</tr>
<tr>
<td>add, mul, [su]div, smod, [su]rem, sub</td>
<td>arithmetic</td>
<td>$B^n \times B^n \rightarrow B^n$</td>
</tr>
<tr>
<td>[su]addo, [su]divo, [su]mulo, [su]subo</td>
<td>overflow</td>
<td>$B^n \times B^n \rightarrow B^1$</td>
</tr>
<tr>
<td>concat</td>
<td>concatenation</td>
<td>$B^n \times B^n \rightarrow B^{n+m}$</td>
</tr>
<tr>
<td>read</td>
<td>array read</td>
<td>$A^{I\rightarrow E} \times I \rightarrow E$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ternary</th>
<th></th>
<th>$B^1 \times B^n \times B^n \rightarrow B^n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ite</td>
<td>conditional</td>
<td></td>
</tr>
<tr>
<td>write</td>
<td>array write</td>
<td>$A^{I\rightarrow E} \times I \times E \rightarrow A^{I\rightarrow E}$</td>
</tr>
</tbody>
</table>

Partially initializing a register can be achieved by applying a bit-mask to an uninitialized register.

Transition functions for both registers and memories are defined with the next keyword. It takes the current and next states as arguments. A state variable without associated next function is treated as a primary input, i.e., it has the same behaviour as inputs defined via keyword input. Note that BTOR provides a next keyword for registers and an anext keyword for memories. Using sorts in BTOR2 avoids such sort specific keyword variants.

As in the latest version of AIGER [4], BTOR2 supports bad state properties, which are essentially negations of safety properties. Multiple properties can be specified by simply adding multiple bad state properties. Invariant constraints can be introduced via the constraint keyword and are assumed to hold globally. A witness for a bad state property is an initialized finite path, which reaches (actually, contains) a bad state and satisfies all invariant constraints.

A witness for a bad state property is an initialized finite path, which reaches (actually, contains) a bad state and satisfies all invariant constraints.

As in AIGER [4], keywords fair and justice allow to specify (global) fairness constraints and (negations of) liveness properties. Each justice property consists of a set of Büchi conditions. A witness for a justice property is an infinite initialized path on which all Büchi conditions and all global fairness constraints
are satisfied infinitely often. In addition, all global invariant constraints have to hold. The justice keyword takes a number (the number of Büchi conditions) and an arbitrary number of nodes (the Büchi conditions) as arguments.

3 Witness Format

The syntax of the BTOR2 witness format is shown in Fig. 2. A BTOR2 witness consists of a sequence of valid input assignments grouped by (time) frames. It starts with 'sat' followed by a list of properties that are satisfied by the witness. A property is identified by a prefix 'b' (for bad) and 'j' (for justice) followed by a number $i$, which ranges over the number of defined bad and justice properties starting from 0. For example, 'b0 j0' refers to the first bad and first justice property in the order they occur in the BTOR2 input. The list of properties is followed by a sequence of $k+1$ frames at time $t \in \{0, \ldots, k\}$. A frame is divided into a state and input part. The state part starts with '#t' and is mandatory for the first frame ($t = 0$) and optional for later frames ($t > 0$). It contains state assignments at time $t$. The input part starts with '@t' and consists of input assignments of the transition from time $t$ to $t + 1$. If states are uninitialized (no init), their initial assignment is required to be specified in frame '#0'. The state part is usually omitted for $t > 0$ since state assignments can be computed from states and inputs at time $t - 1$. While don’t care inputs can be omitted, our witness checker assumes that they are zero. Input and state assignments use the same numbering scheme as properties, i.e., states and inputs are numbered separately in the order they are defined, starting from 0. For example, 0 in frame '#t' (or '@t') refers to the first state (or input) as defined in the BTOR2 input. For justice properties we assume the witness to be lasso shaped, i.e., the next state, which can be computed from the last state and inputs at time $k$, is identical to one of the previous states at time $t = 0 \ldots k$. As in AIGER, a BTOR2 witness is terminated with '.' on a separate line.

\[
\text{(binary-string)} ::= [0-1]+ \\
\text{(bv-assignment)} ::= (\text{binary-string}) \\
\text{(array-assignment)} ::= \left[ (\text{binary-string}) \right] (\text{binary-string}) \\
\text{(assignment)} ::= (\text{uint}) (\text{bv-assignment} \mid (\text{array-assignment})) (\text{symbol}) \\
\text{(model)} ::= (\text{uint}) (\text{bv-assignment} \mid (\text{array-assignment})) (\text{symbol}) \\
\text{(state part)} ::= (\text{uint}) (\text{symbol}) \\
\text{(input part)} ::= (\text{uint}) (\text{symbol}) \\
\text{(frame)} ::= (\text{frame}) (\text{input part}) \\
\text{(prop)} ::= (\text{uint}) (\text{prop}) (\text{symbol}) \\
\text{(header)} ::= \text{sat} (\text{prop}) (\text{symbol}) \\
\text{(witness)} ::= (\text{frame}) (\text{symbol}) (\text{prop}) (\text{symbol}) (\text{symbol})
\]

Fig. 2. BTOR2 model and witness format syntax (sequential part in red). (Color figure online)

Figure 3 illustrates a simple C program (left), the corresponding BTOR2 model with the negation of the assertion as a bad property (center), and a
BTOR2 witness for the violated property (right). The BTOR2 model defines one bad property \((a == 3 \&\& b == 3)\), which is satisfied in frame 6. The corresponding witness identifies this property as bad property 'b0' (first bad property defined in the model). All states are initialized, hence ‘#0’ is empty, and ‘@0’ to ‘@6’ indicate the assignments of input 0 (turn, the first input defined in the model) in frames 0 to 6, e.g., turn = 1 at t = 0, turn = 0 at t = 1 and so on. In frame 6, both states a and b reach value 3, and therefore property 'b0' is satisfied.

4 Tools

We provide a generic stand-alone parser for BTOR2, which features basic type checking and consists of approx. 1,500 lines of C code. We implemented a reference bounded model checker BtorMC, which currently supports checking safety (aka. bad state) properties for models with registers and memories and produces witnesses for satisfiable properties. Unrolling the model is performed by symbolic simulation, i.e., symbolic substitution of current state expressions into next state functions, and incremental SMT solving. We also implemented a simulator for randomly simulating BTOR2 models. It further supports checking BTOR2 witnesses. The model checker is tightly integrated into our SMT solver Boolector [18], an award-winning SMT solver for the theory of fixed-size bit-vectors with arrays and uninterpreted functions. Since the last major version [18], we extended Boolector with several new features. Most notably, Boolector 3.0 now comes with support for quantified bit-vectors [24] and two different local search strategies for quantifier-free bit-vector formulas that don’t rely on but can be combined with bit-blasting [19,21,22]. It further provides support for BTOR2. In contrast to previous versions of Boolector, Boolector 3.0 and all BTOR2 tools.
are released under the MIT open source license and the source code is hosted on GitHub\(^1\).

## 5 Experiments

We collected ten real-world (System)Verilog designs with safety properties from various open source projects [11,26–28]. The majority of these designs include memories. We used the open synthesis suite Yosys [29] to synthesize these designs into BTOR2 and SMT-LIB. For BTOR2, Yosys directly generates the models from a circuit description. For SMT-LIB, since the language does not support describing model checking problems, we used Yosys in combination with Yosys-SMTBMC to produce unrolled (incremental) problems.

We compared BtorMC against the most recent versions of Boolector (3.0) and Yices [10] (2.5.4), the two best solvers of the QF_ABV division of the SMT competition 2017. The BTOR2 models serve as input for BtorMC, and the incremental SMT-LIB benchmarks serve as input for Boolector and Yices. All benchmarks, synthesis scripts, generated files, log files and the source code of our tools for this evaluation are available at http://fmv.jku.at/cav18-btor2.

The results in Table 2 show that our flow using BTOR2 as intermediate format is competitive with simple unrolling. Note that our model checker BtorMC issues incremental calls to Boolector. However, in Boolector, sophisticated word-level rewriting is currently disabled in incremental mode. We expect a major performance boost by fully supporting incremental word-level preprocessing.

**Table 2.** BtorMC/BTOR2 vs. unrolled SMT-LIB with a time limit of 3600 s, where \(k\) is the bound and \#bad is the number of bad properties.

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>(k)</th>
<th>#bad</th>
<th>BtorMC time[s]</th>
<th>Boolector time[s]</th>
<th>Yices time[s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>picorv32-check</td>
<td>30</td>
<td>23</td>
<td><strong>4.8</strong></td>
<td>18.9</td>
<td>10.8</td>
</tr>
<tr>
<td>picorv32-pcregs</td>
<td>20</td>
<td>3</td>
<td><strong>63.0</strong></td>
<td>293.0</td>
<td>TO</td>
</tr>
<tr>
<td>ponylink-slaveTXlen-sat</td>
<td>230</td>
<td>1</td>
<td>305.5</td>
<td>406.8</td>
<td><strong>145.6</strong></td>
</tr>
<tr>
<td>ponylink-slaveTXlen-unsat</td>
<td>231</td>
<td>1</td>
<td>183.8</td>
<td>131.4</td>
<td><strong>71.4</strong></td>
</tr>
<tr>
<td>VexRiscv-regch0-15</td>
<td>17</td>
<td>2</td>
<td><strong>9.6</strong></td>
<td>48.3</td>
<td>12.2</td>
</tr>
<tr>
<td>VexRiscv-regch0-20</td>
<td>22</td>
<td>2</td>
<td>528.8</td>
<td><strong>520.7</strong></td>
<td>2232.2</td>
</tr>
<tr>
<td>VexRiscv-regch0-30</td>
<td>32</td>
<td>2</td>
<td>TO</td>
<td>TO</td>
<td>TO</td>
</tr>
<tr>
<td>zipcpu-busdelay</td>
<td>100</td>
<td>50</td>
<td><strong>157.0</strong></td>
<td>287.0</td>
<td>181.2</td>
</tr>
<tr>
<td>zipcpu-pfcache</td>
<td>100</td>
<td>39</td>
<td><strong>17.4</strong></td>
<td>19.9</td>
<td>32.5</td>
</tr>
<tr>
<td>zipcpu-zipmmmu</td>
<td>30</td>
<td>57</td>
<td>86.0</td>
<td>412.9</td>
<td><strong>46.5</strong></td>
</tr>
</tbody>
</table>

\(^1\) https://github.com/boolector.
6 Conclusion

We propose BTOR2, a new word-level model-checking and witness format. For this format we provide a generic parser implementation, a simulator that also checks witnesses, and a reference bounded model checker BtorMC, which is tightly integrated with our SMT solver Boolector. These open source tools are evaluated on new real-world benchmarks, which we synthesized from open source hardware (System) Verilog models into BTOR2 and SMT-LIB with Yosys. The tool Verilog2SMV [14] translates Verilog into model-checking problems in several formats, including nuXmv [7] and BTOR. However, its translation to BTOR is incomplete and development discontinued.

We plan to provide a translator from BTOR2 into SALLY [25], and VMT [8], which are both extensions of SMT-LIB to model symbolic transition systems. It might also be interesting to translate incremental SMT-LIB benchmarks and horn clause models (as handled by, e.g., µZ [13]) into BTOR2 and vice versa. We hope other compilers and model checkers such as SAL [9], EBMC [15] and ABC [12, 16] will provide support to produce and read BTOR2 models. We want to extend the format to other logics, in particular to support lambdas as in [23]. There is also a need for fuzzing [20] and delta-debugging tools [17].

Last but not least, we want to use this format to bootstrap a word-level model checking competition, which of course needs more benchmarks.

References


Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.