

# Causality-based Model Checking \*

Bernd Finkbeiner

Universität des Saarlandes  
Saarbrücken, Germany

finkbeiner@cs.uni-saarland.de

Andrey Kupriyanov

Institute of Science and Technology Austria  
Klosterneuburg, Austria

andrey.kupriyanov@ist.ac.at

Model checking is usually based on a comprehensive traversal of the state space. Causality-based model checking is a radically different approach that instead analyzes the cause-effect relationships in a program. We give an overview on a new class of model checking algorithms that capture the causal relationships in a special data structure called concurrent traces. Concurrent traces identify key events in an execution history and link them through their cause-effect relationships. The model checker builds a tableau of concurrent traces, where the case splits represent different causal explanations of a hypothetical error. Causality-based model checking has been implemented in the ARCTOR tool, and applied to previously intractable multi-threaded benchmarks.

## 1 Introduction

Model checking [1] is a “push-button” technology: it verifies a given program completely automatically, without any help from the human user. A fundamental challenge, however, is the infamous state space explosion problem: in a concurrent program, the number of control states grows exponentially with the number of parallel threads. This creates a barrier for the standard approach to model checking, which is based on a comprehensive traversal of the state space.

Causality-based model checking [4, 5, 3] is inspired by the observation that it is not unusual for programs that are difficult to verify with a conventional model checker to have surprisingly short “paper-and-pencil” proofs. A likely explanation is that humans reason in terms of *causal relationships*. In most general terms, causality can be defined as the relation between two events, where the first event (the *cause*) is understood to be partly responsible for the second (the *effect*). Reasoning by causality is the usual style of constructing proofs: assume some situation (the effect) to be present, and derive all possible explanations (the causes). Consider the following assertion and its proof from Leslie Lamport’s paper [6] introducing the Bakery algorithm for mutual exclusion:

**Assertion 1.** *If processors  $i$  and  $k$  are in the bakery and  $i$  entered the bakery before  $k$  entered the doorway, then  $number[i] < number[k]$ .*

*Proof.* By hypothesis,  $number[i]$  had its current value while  $k$  was choosing the current value of  $number[k]$ . Hence,  $k$  must have chosen  $number[k] \geq 1 + number[i]$ .  $\square$

The proof starts by assuming the situation where the event “ $i$  entered the bakery” precedes the event “ $k$  entered the doorway”, and, moreover,  $number[i]$  preserves its value between two events. The proof proceeds by deriving from this situation another *necessary* fact (notice the words “*must have chosen*”): ( $number[k] \geq 1 + number[i]$ ).

---

\*This research was supported in part by the European Research Council (ERC) Grant OSARES (No. 683300), by the German Research Foundation (DFG) as part of SFB/TR 14 AVACS, and by the Austrian Science Fund (FWF) under grants S11402-N23 (RiSE and SHiNE) and Z211-N23 (Wittgenstein Award).

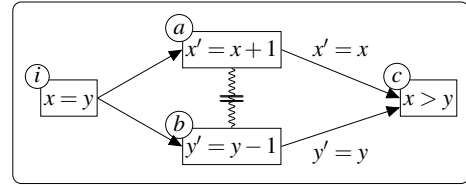
Unlike standard model checking, causality-based model checking is not based on a traversal of the state space but instead tracks the causal dependencies in the system. In concurrent programs, it is often the case that not many concurrent events depend on each other – most events are, in fact, independent, and precisely this allows concurrent programs to achieve better performance than sequential programs. Causality-based model checking provides a formal proof system as well as an automatic method for constructing proofs or finding counterexamples employing the principles of causal reasoning.

## 2 The Causality-based Model Checking Framework

We now introduce the main components of the causality-based model checking framework.

**Concurrent traces.** The basic building block, which we intend as a replacement of *state* in the standard model checking, is called a *concurrent trace*. Instead of a single momentary snapshot of the program computation, it represents a set of related computation events. Each event is labeled by a *transition predicate*, and describes a set of program transitions satisfying the predicate. Events are related by *causal links*, which are simply the ordering constraints between events. Causal links are also labeled by transition predicates; these constraints represent conditions which all events in scope of the causal link should satisfy. Finally, there is a *conflict relation*, which prohibits certain events to coincide in time.

A concurrent trace as a whole can be understood as a combination of existential and universal constraints, which collectively describe a set of program computations, such that the concurrent trace can be mapped to each of those computations respecting the constraints. E.g., a concurrent trace on the right represents the set of computations which satisfy the following formula:



$$\begin{aligned} \exists i, a, b, c. & (i \leq a) \wedge (i \leq b) \wedge (a \leq c) \wedge (b \leq c) \wedge (a \neq b) \wedge \\ & (s_i, s_{i+1}) \models (x = y) \wedge (s_a, s_{a+1}) \models (x' = x + 1) \wedge (s_b, s_{b+1}) \models (y' = y - 1) \wedge (s_c, s_{c+1}) \models (x > y) \\ \wedge \forall j, k. & (a < j < c) \implies (s_j, s_{j+1}) \models (x' = x) \wedge (b < k < c) \implies (s_k, s_{k+1}) \models (y' = y) \end{aligned}$$

In [5] we have extended the basic model of *finite* concurrent traces from [4] to *infinite* concurrent traces. An infinite concurrent trace consists of a *stem* and a *cycle*, each being a concurrent trace, with the semantics that the cycle occurs infinitely often after the stem occurs once.

*Graphical Notation.* We show event identities in circles, and labeling formulas in squares. Causal links are shown as solid lines with arrows, and conflicts as crossed zigzag lines. The cycle part of the trace is depicted in round brackets, superscripted with  $\omega$ . We omit any of those when they are not important or would create clutter in the current context.

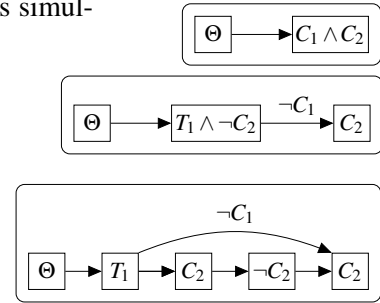
**Representation of program properties.** Given a program property of interest, we start our analysis by representing the *property violation* as a concurrent trace. From our experience, we find out that violations of many useful properties can be quite naturally expressed as concurrent traces.

Suppose we are given a program  $P$ , consisting of two processes  $P_1$  and  $P_2$  that require mutually exclusive access to their critical sections. Each process  $P_i$ ,  $i = 1, 2$ , contains an endless loop  $\langle N_i \rightarrow T_i \rightarrow C_i \rangle$  with three sections:  $N_i$  (“noncritical”),  $T_i$  (“trying”), and  $C_i$  (“critical”). We can easily represent violations of the desired properties of  $P$  as concurrent traces (we denote the initial state of  $P$  as  $\Theta$ ).

*Mutual exclusion.* Processes should not be in their critical sections simultaneously, which is described by the formula  $\Box \neg(C_1 \wedge C_2)$ .

*Strict precedence.* When  $P_1$  is trying to get access and  $P_2$  is not in the critical section, then  $P_1$  will be admitted to the critical section first:  $\Box((T_1 \wedge \neg C_2) \implies (\neg C_2) \not\sim C_1)$ .

*Bounded overtaking.* We may want to give an upper bound on the amount of overtaking, where overtaking means that one process enters the critical section ahead of its rival. The property of 1-bounded overtaking can be expressed as  $\Box(T_1 \implies (\neg C_2) \not\sim (C_2) \not\sim (\neg C_2) \not\sim (C_1))$ .



**Trace transformers.** Our analysis proceeds in steps, where each step takes some concurrent trace, and applies to it a so called *trace transformer*. Trace transformers are the proof rules describing necessary consequences from the current analysis situation represented as a concurrent trace. There is a number of general proof rules for safety and liveness analysis that were presented in our previous work [3, 4, 5]; we also envision that application-specific proof rules can be easily created for particular domains.

A trace transformer is an ordered set of *trace productions* of the form  $(L \xrightarrow{r_i} R_i)$ , where all productions share the same left-hand side  $L$ . A trace production is a generalization of a *graph production* [7] and describes a formal rule to transform one concurrent trace into another. Collectively, all productions of a trace transformer encode some case distinction, where the union of trace languages of individual cases is contained in the language of the original trace. Below we show several examples of trace transformers.

*OrderSplit(a, b)* trace transformer considers alternative orderings of two concurrent events  $a$  and  $b$ .

*NecessaryEvent(a, b, φ)*, given two causally related and conflicting events  $a$  and  $b$  in a concurrent trace, and a state predicate  $\varphi \in \Phi(\mathcal{V})$ , such that the label of  $a$  implies  $\varphi'$ , and the label of  $b$  implies  $\neg\varphi$ , introduces a new “bridging” event  $c$  in between. This condition can be interpreted

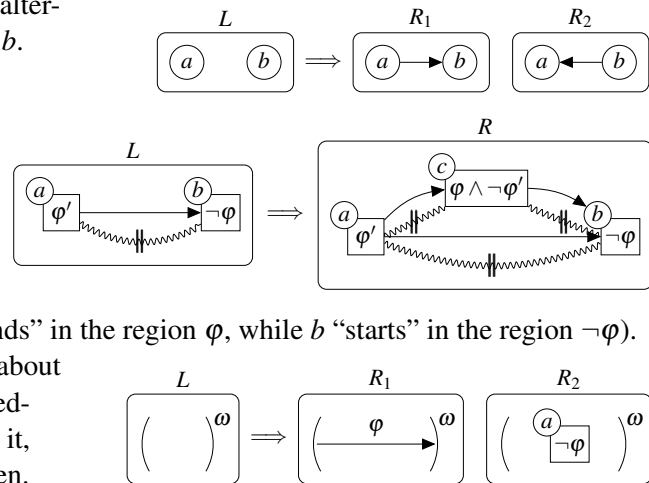
as a contradiction between events  $a$  and  $b$  ( $a$  “ends” in the region  $\varphi$ , while  $b$  “starts” in the region  $\neg\varphi$ ).

*InvarianceSplit(φ)* makes a case distinction about the program behavior at infinity: for a given predicate  $\varphi$  either all events in the cycle part satisfy it, or a violating event should happen infinitely often.

Due to the space limitations we are not able to show here, even informally, all trace transformers used in the examples of the next section; the interested reader should refer to [3, 4, 5] for their formal description.

**Tableau construction.** The final component of our framework is a collection of datastructures for the organization of proof or counterexample search. Each datastructure comes with the corresponding algorithm for its automatic construction; the datastructures and algorithms have different properties which are detailed in Figure 1.

The conceptually simplest datastructure, called *trace unwinding*, is a forest of nodes each labeled with a concurrent trace. The forest roots represent concurrent traces that encode possible violations of



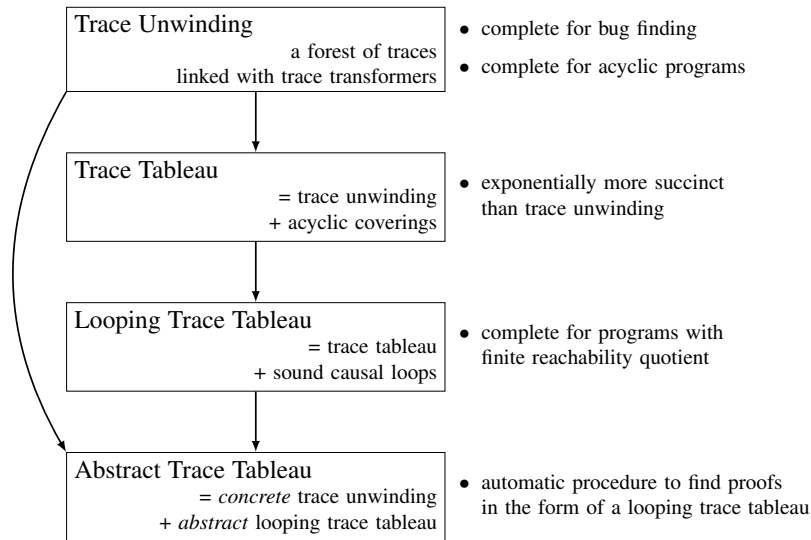


Figure 1: Data structures for causality-based model checking.

the property of interest. The exploration algorithm proceeds by picking some forest leaf, and employing an applicable trace transformer, producing a number of further nodes. The exploration stops when all forest leaves are found to be contradictory.

The most involved datastructure, *abstract trace tableau*, contains, besides concrete trace unwinding, an abstract *looping trace tableau*, which may have covering edges between tableau nodes. A covering condition is an extension of subgraph isomorphism to concurrent traces: a concurrent trace, which is a subgraph of another trace, represents a situation which was encountered in the analysis before. This tableau is also allowed to contain *causal loops*, i.e. infinite repetitions of a sequence of trace productions, which together imply the impossibility of a computation satisfying them. The abstract trace tableau is used to track premises of already applied proof rules, and, thus, simplifies coverings.

### 3 Two Examples

We illustrate the framework with two examples taken from [4] and [5]: one for the analysis of a safety property (reachability), and another for the analysis of a liveness property (termination).

**Safety.** Consider the synchronized system shown in the top part of Figure 2: the example was introduced by Esparza and Heljanko in [2] to illustrate the exponential succinctness of Petri net unfoldings. There are  $n + 1$  processes, and we want to check whether the global transition  $\mathbf{c}$  is executable. Note that the state space of this system is exponential with respect to  $n$ : the system contains  $3 \cdot 2^{n-1}$  reachable states. Thus, approaches based on state space exploration will suffer from the state space explosion problem. The authors of [2] show that the Petri net unfolding of the example system contains  $2 \cdot n + 3$  places, i.e., a *linear size* unfolding can represent succinctly the exponential state space.

We use the same example to demonstrate that the trace unwinding of the example system never exceeds  $n + 6$  nodes, but a *constant size* unwinding of just 7 nodes also suffice, which we show in the bottom part of Figure 2. Node 1, the root of the unwinding, captures all system traces where  $\mathbf{c}$  is executed. One of its preconditions is that the first process should be at location  $r_2$ ; but the initial condition says that the system is at location  $r_1$ : a contradiction. Thus, a transition that goes from  $r_1$  to  $r_2$  is *necessary*,

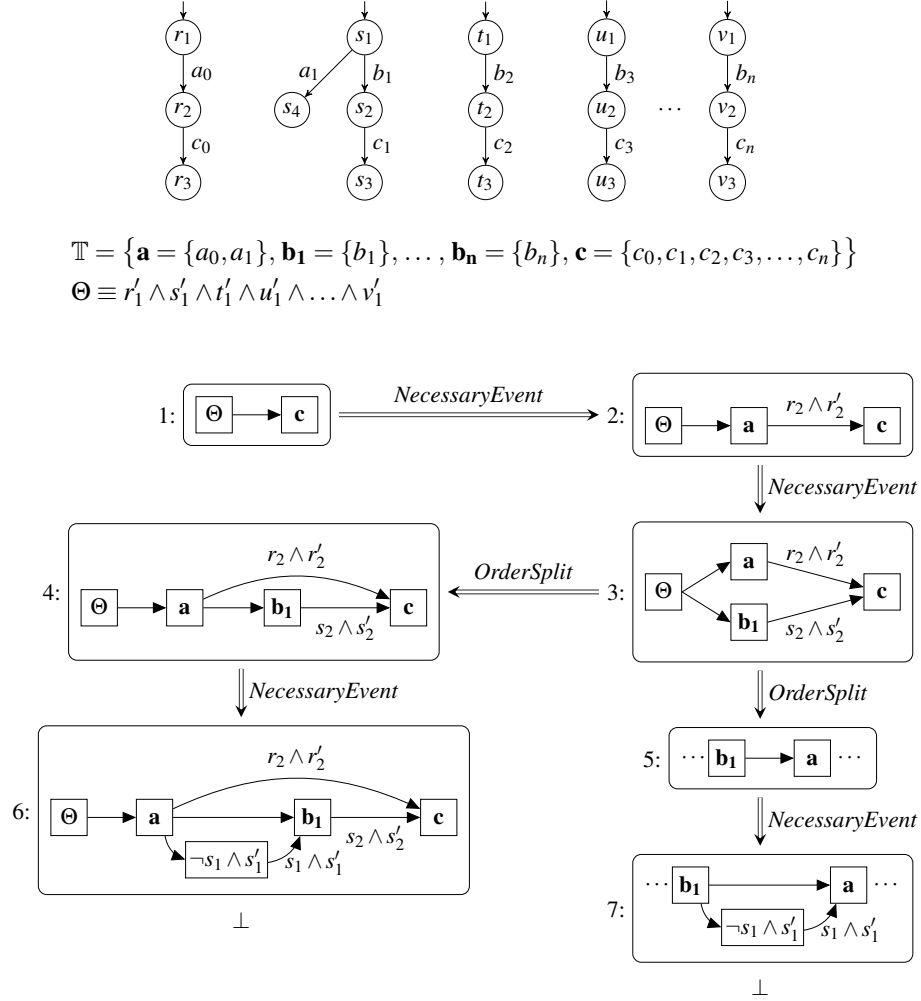


Figure 2: *Top*: example system from [2]. *Bottom*: its correctness proof in the form of a trace unwinding.

and we insert the only such transition,  $\mathbf{a}$ , into the trace of node 2. Formally, this is done by applying the trace transformer *LastNecessaryEvent*, this is where the link label  $r_2 \wedge r'_2$  comes from: it enforces to select the last transition that goes into location  $r_2$ . By a similar reasoning we conclude that transition  $\mathbf{b}_1$  is also necessary, and include it into the trace of node 3. Notice that these events occur concurrently, i.e. no specific order between them is specified till now.

In the next iteration, when we try to put them in some linear order, we find out that these transitions contradict each other:  $\mathbf{a}$  goes to location  $s_4$ ,  $\mathbf{b}_1$  goes to location  $s_2$ , but both can start only at location  $s_1$ . Therefore, we perform a case split between two possible linearizations using the transformer *OrderSplit*, and obtain the traces of nodes 4 and 5. In node 4 the contradiction between now linearly ordered transitions  $\mathbf{a}$  and  $\mathbf{b}_1$  is enforced by the trace, and we again apply the transformer *LastNecessaryEvent*. It inserts an event between transitions  $\mathbf{a}$  and  $\mathbf{b}_1$  which needs to change the location from  $s_4$  (the postcondition of  $\mathbf{a}$ ) to  $s_1$  (the precondition of  $\mathbf{b}_1$ ). Formally, this requirement is captured by the Craig interpolant between the post- and pre-conditions of  $\mathbf{a}$  and  $\mathbf{b}_1$ , respectively, which happens to be  $\neg s_1$ . As there is no transition that goes from a location different from  $s_1$  to  $s_1$ , this trace is declared as contradictory, and the

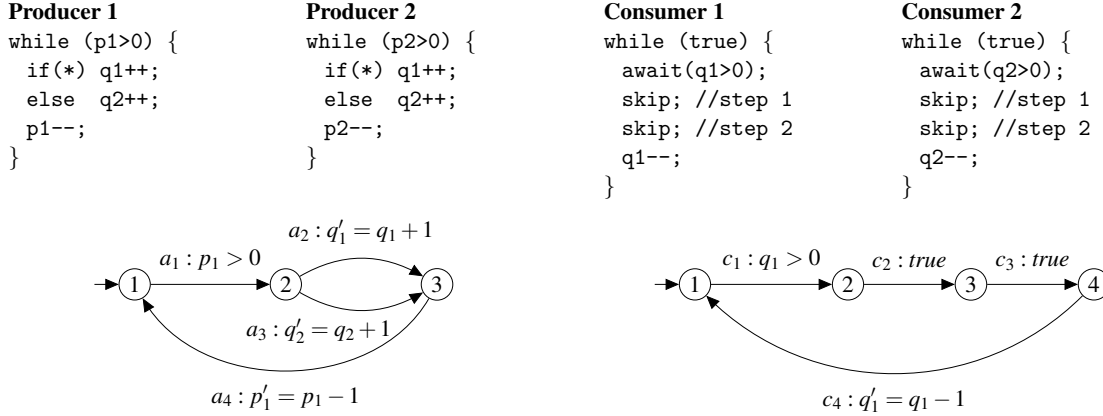


Figure 3: The *Producer-Consumer* benchmark, shown here for 2 producers and 2 consumers (*Top*: pseudocode; *Bottom*: control flow graphs with labeled transitions for Producer 1 and Consumer 1). The producer threads draw tasks from individual pools and distribute them to nondeterministically chosen queues, each served by a dedicated consumer thread; two steps are needed to process a task. The integer variables  $p_1$  and  $p_2$  model the number of tasks left in the pools of Producers 1 and 2, the integer variables  $q_1$  and  $q_2$  model the number of tasks in the queues of Consumers 1 and 2.

left branch of the unwinding is closed. For the right branch of the unwinding, consisting of nodes 5 and 7, we proceed in the same way, and also close it as contradictory. Thus, the unwinding is complete, and we conclude that transition **c** is not executable.

The unwinding of Figure 2 has constant size, independent on the number of processes in the example system. In the worst case it could have linear size: for that to happen, the other necessary transitions **b**<sub>2</sub> through **b**<sub>n</sub> would have to be introduced into concurrent before transitions **a** and **b**<sub>1</sub> are introduced. In fact, a simple heuristic is able to select transitions **a** and **b**<sub>1</sub> first, and it will always produce the trace unwinding of constant size.

**Liveness.** Consider the Producer-Consumer example presented in Figure 3, which is a simplified model of the *Map-Reduce* architecture from distributed processing: producers model the mapping step for separate data sources, consumers model the reducing step for different types of input data. The natural requirement for this architecture is that the processing terminates for any finite amount of input data.

Our analysis starts with the assumption (by way of contradiction) that there exists some infinite computation. The assumption is expressed as the concurrent trace of node 1 in Figure 4: infinitely often some transition should occur. The transition is so far unknown, and therefore characterized by the predicate  $\top$ . Our argument proceeds by instantiating this unknown transition with the transitions of the program, resulting in one new trace per transition.

For example, transition  $a_1$  of Producer 1, gives us the trace of node 2. The consequence of the decision that  $a_1$  occurs infinitely often is that  $a_4$  must also occur infinitely often:  $a_1$  starts at location 1, and only  $a_4$  can bring Producer 1 to this location. The requirement that both  $a_1$  and  $a_4$  occur infinitely often is expressed as the trace of node 3, obtained from the trace of node 1 by the *NecessaryCycleEvent* trace transformer. The trace of node 3 is terminating:  $p_1$  is decreased infinitely often and is bounded from below; it is therefore a ranking function. An infinite computation might exist only if some transition increases  $p_1$  (satisfies the predicate  $p_1' > p_1$ ) is executed infinitely often. This situation is expressed by the trace of node 4, obtained by the application of the *InvarianceSplit* trace transformer. Since there is no transition in the program transition relation that satisfies  $p_1' > p_1$ , we arrive at a contradiction.

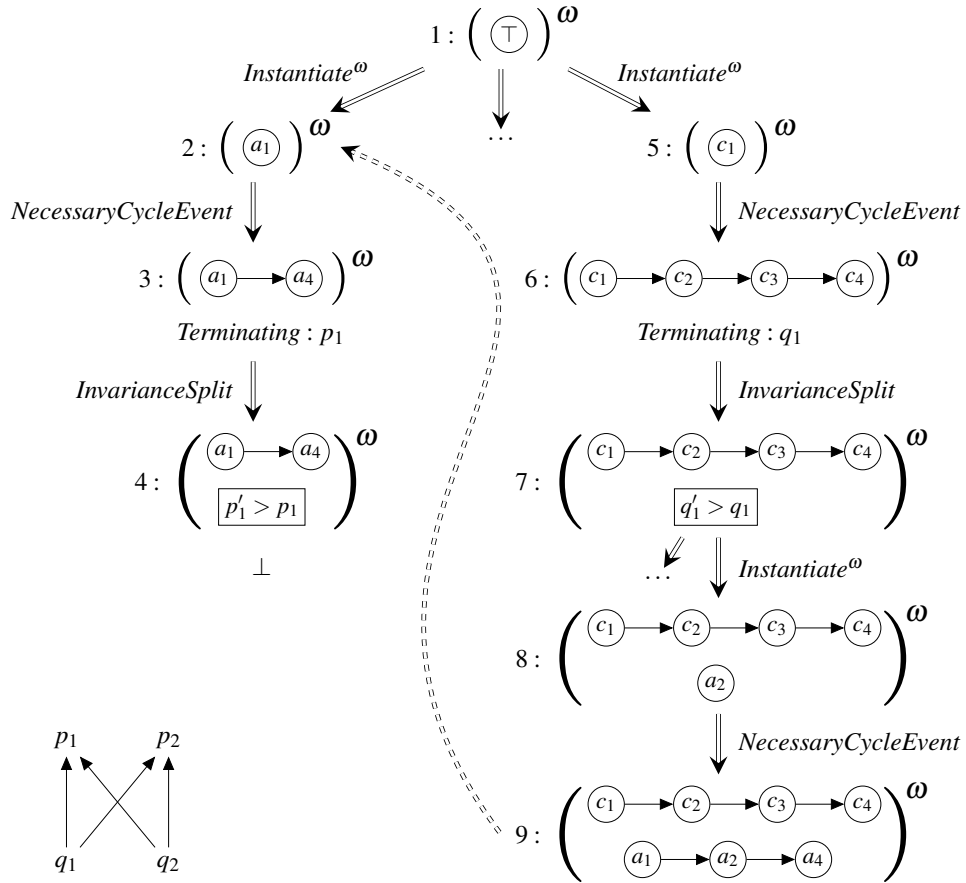


Figure 4: Termination proof for the Producer-Consumer example of Figure 3. *Bottom left*: partially ordered ranking function discovered in the analysis.

Let us explore another instantiation of the unknown event in the trace of node 1, this time with transition  $c_1$  of Consumer 1: we obtain the trace of node 5. Again, exploring causal consequences, local safety analysis gives us that events  $c_2$ ,  $c_3$ , and  $c_4$  should also occur infinitely often in the trace: we insert them, and get the trace of node 6. Termination analysis for that trace gives us the ranking function  $q_1$ : it is bounded from below by event  $c_1$  and decreased by event  $c_4$ . Again, we conclude that the event increasing  $q_1$  should occur infinitely often, and introduce it in the trace of node 7.

Next, we try all possible instantiations of the event characterized by the predicate ( $q'_1 > q_1$ ): there are two transitions that satisfy the predicate, namely  $a_2$  and  $b_2$ . We explore the instantiation with  $a_2$  in the trace of node 8, and see that transitions  $a_1$  and  $a_4$  should occur infinitely often (node 9). At this point, we realize that the trace of node 9 contains as a subgraph the trace of node 2, namely the transition  $a_1$ . We cover node 9 with node 2, and avoid repeating the analysis done for nodes 2–4. The remaining tableau branches are analyzed similarly. The resulting tableau for the case of two producers and two consumers will have the shape shown in the bottom left part of Figure 4. It can be interpreted as a partially ordered ranking function, which shows that all threads satisfying the function components  $p_1$  and  $p_2$  terminate unconditionally, while the threads that satisfy the function components  $q_1$  and  $q_2$  terminate under the condition that both of the previous components terminate. Notice also that the tableau is of quadratic size with respect to the number of threads.

Threads	Terminator		T2		AProVE		ARCTOR		
	Time(s)	Mem.(MB)	Time(s)	Mem.(MB)	Time(s)	Mem.(MB)	Time(s)	Mem.(MB)	Vertices
1	3.37	26	2.42	38	3.17	237	0.002	2.3	6
2	1397	1394	3.25	44	6.79	523	0.002	2.6	11
3	×	MO	U(29.2)	253	U(26.6)	1439	0.002	2.6	21
4	×	MO	U(36.6)	316	U(71.2)	1455	0.003	2.7	30
5	×	MO	U(30.7)	400	U(312)	1536	0.007	2.7	44
20	×	MO	Z3-TO	×	×	MO	0.30	4.2	470
40	×	MO	Z3-TO	×	×	MO	4.30	12.7	1740
60	×	MO	Z3-TO	×	×	MO	20.8	35	3810
80	×	MO	Z3-TO	×	×	MO	67.7	145	6680
100	×	MO	Z3-TO	×	×	MO	172	231	10350

Table 1: Running times of the termination provers *Terminator*, *T2*, *AProVE*, and *ARCTOR* on the Producer-Consumer benchmark [5]. MO stands for memout. U indicates that the termination prover returned “unknown”; Z3-TO indicates a timeout in the Z3 SMT solver.

## 4 Conclusions

Causality-based model checking has significant advantages over standard state-based model checking. As illustrated by the two examples, the complexity of the verification problem can be substantially lower; in particular, the complexity of verifying multi-threaded programs with locks reduces from exponential to polynomial (see [4], Theorem 4). The efficiency of causality-based model checking is also reflected in the experimental results obtained with our tool implementation *ARCTOR* (cf. [5]). In our experience, *ARCTOR* scales to many, even hundreds, of parallel threads in benchmarks where other tools can only handle a small number of parallel threads or no parallelism at all: see Table 1 for experimental results obtained with *ARCTOR* on the Producer-Consumer example.

## References

- [1] Edmund M. Clarke & E. Allen Emerson (1981): *Design and Synthesis of Synchronization Skeletons Using Branching-Time Temporal Logic*. In: *Logics of Programs, Workshop, Yorktown Heights, New York, May 1981*, pp. 52–71, doi:10.1007/BFb0025774.
- [2] Javier Esparza & Keijo Heljanko (2008): *Unfoldings - A Partial-Order Approach to Model Checking*. Monographs in Theoretical Computer Science. An EATCS Series, Springer, doi:10.1007/978-3-540-77426-6.
- [3] Andrey Kupriyanov (2016): *Causality-based verification*. Ph.D. thesis, Saarland University, Saarbrücken, Germany. Available at <http://scidok.sulb.uni-saarland.de/volltexte/2016/6696/>.
- [4] Andrey Kupriyanov & Bernd Finkbeiner (2013): *Causality-Based Verification of Multi-threaded Programs*. In: *CONCUR 2013 - Concurrency Theory - 24th International Conference, CONCUR 2013, Buenos Aires, Argentina, August 27-30, 2013. Proceedings*, pp. 257–272, doi:10.1007/978-3-642-40184-8\_19.
- [5] Andrey Kupriyanov & Bernd Finkbeiner (2014): *Causal Termination of Multi-threaded Programs*. In: *Computer Aided Verification - 26th International Conference, CAV 2014, Vienna, Austria, July 18-22, 2014. Proceedings*, pp. 814–830, doi:10.1007/978-3-319-08867-9\_54.
- [6] Leslie Lamport (1974): *A New Solution of Dijkstra’s Concurrent Programming Problem*. *Commun. ACM* 17(8), pp. 453–455, doi:10.1145/361082.361093.
- [7] Grzegorz Rozenberg, editor (1997): *Handbook of Graph Grammars and Computing by Graph Transformations, Volume 1: Foundations*. World Scientific.