Verification of Large State/Event Systems Using Compositionality and Dependency Analysis*

Jørn Lind-Nielsen[1], Henrik Reif Andersen[1], Gerd Behrmann[2],
Henrik Hulgaard[1], Kåre Kristoffersen[2], and Kim G. Larsen[2]

[1] Department of Information Technology, Technical University of Denmark
e-mail: {jl,hra,henrik}@it.dtu.dk

[2] BRICS**, Department of Computer Science, Aalborg University
e-mail: {behrmann,jelling,kgl}@cs.auc.dk

Abstract. A state/event model is a concurrent version of Mealy machines used for describing embedded reactive systems. This paper introduces a technique that uses *compositionality* and *dependency analysis* to significantly improve the efficiency of symbolic model checking of state/event models. This technique makes possible automated verification of large industrial designs with the use of only modest resources (less than one hour on a standard PC for a model with 1421 concurrent machines). The results of the paper are being implemented in the next version of the commercial tool visualSTATE™.

1 Introduction

Symbolic model checking is a powerful technique for formal verification of finite-state concurrent systems. The technique has proven very efficient for verifying hardware systems: circuits with an extremely large number of reachable states has been verified. However, it is not clear whether model checking is effective for other kinds of concurrent systems as, for example, software systems. One reason that symbolic model checking may not be as efficient is that software systems tend to be both larger and less regularly structured than hardware. For example, many of the results reported for verifying large hardware systems have been for linear structures like stacks or pipelines (see, e.g., [8]) for which it is known [19] that the size of the transition relation (when represented as an ROBDD) grows linearly with the size of the system. Only recently has the first experiments on larger realistic software systems been reported [3, 20].

This paper presents a technique that significantly improves the performance of symbolic model checking on embedded reactive systems modeled using a

---

state/event model. The state/event model is a concurrent version of Mealy machines, that is, it consists of a fixed number of concurrent finite state machines that have pairs of input events and output actions associated with the transitions of the machines. The model is synchronous: each input event is reacted upon by all machines in lock-step; the total output is the multi-set union of the output actions of the individual machines. Further synchronization between the machines is achieved by associating a guard with the transitions. Guards are Boolean combinations of conditions on the local states of the other machines. In this way, the firing of transitions in one machine can be made conditional on the local state of other machines. If a machine has no enabled transition for a particular input event, it simply does not perform any state change.

The state/event model is convenient for describing the control portion of embedded reactive systems, including smaller systems as cellular phones, hi-fi equipment, and cruise controls for cars, and large systems as train simulators, flight control systems, telephone and communication protocols. The model is used in the commercial tool visualSTATE™ [18]. This tool assists in developing embedded reactive software by allowing the designer to construct a state/event model and analyze it by either simulating it or by running a *consistency checker.* The tool automatically generates the code for the hardware of the embedded system. The consistency checker is in fact a verification tool that checks for a range of properties that any state/event model should have. Some of the checks must be passed for the generated code to be correct, for instance, it is crucial that the model is deterministic. Other checks are issued as warnings that might be design errors such as transitions that can never fire.

State/event models can be extremely large. And unlike in traditional model checking, the number of checks is at least linear in the size of the model. This paper reports results for models with up to 1421 concurrent state machines (10^{476} states). For systems of this size, traditional symbolic model checking techniques fail, even when using a partitioned transition relation [6] and backwards iteration. We present a *compositional* technique that initially considers only a few machines in determining satisfaction of the verification task and, if necessary, gradually increases the number of considered machines. The machines considered are determined using a *dependency analysis* of the structure of the system.

The results are encouraging. A number of large state/event models from industrial applications have been verified. Even the largest model with 1421 concurrent machines can be verified with modest resources (it takes less than an hour on a standard PC). Compared with the current version of visualSTATE™, the results improve on the efficiency of checking the smaller instances and dramatically increase the size of systems that can be verified.

Related Work

The use of ROBDDs [5] in model checking was introduced by Burch *et al.* [7] and Coudert *et al.* [13]. Several improvements have been developed since, such as using a partitioned transition relation [6,14] and simplifying the ROBDD representation during the fixed-point iteration [12]. Many of these improvements

are implemented in the tool SMV [19]. Other techniques like abstraction [10] and compositional model checking [11] further reduce the complexity of the verification task, but require human insight and interaction.

Several techniques attempt to improve the efficiency of the verification by only considering a subset of the system. For example, [4] presents a conservative technique for showing emptiness of L-processes based on including only a subset of the processes. The technique is based on analyzing an error trace from the verifier, and use this trace to modify the considered subset of L-processes.

A more promising technique, based on ROBDDs, that also exploits the structure of the system is presented in [17]. On the surface this technique is very close to the one presented here and thus we will discuss it in more detail. The technique in [17] uses a partitioned transition relation and a greedy heuristic is used to select subsets of the transition relation. For each chosen subset, a complete fixed-point iteration is performed. If the formula cannot be proven after this iteration, a larger subset is chosen. In case of an invalid formula the algorithm only terminates when the full transition relation has been constructed (or memory or time has been exhausted). To compare this approach with the one presented here, we can consider a subset of the transition relation as being similar to a subset of the machines in the state/event model. The approach of [17] differs from ours in several central aspects:

- In selecting a new machine to include in the transition relation, [17] uses a greedy strategy involving a fixed-point iteration for each of the remaining machines. (If the system only has a single initial state—as in state/event systems—the greedy strategy reduces to selecting an arbitrary machine.) We chose a new machine based on an initial dependency analysis and thus avoids any extraneous fixed-point iterations.
- Due to a central monotonicity result (lemma 1), we can reuse the previously computed portion of the state space instead of having to start from scratch each time a new machine is added.
- In case the property to be verified is invalid, we only include those machines that are actually dependent on each other in the transition relation. In these cases, [17] may have to include all machines to disprove the property.
- Even when all machines are needed, experiments have shown that our technique of including machines one at a time (exploiting the monotonicity property) is *faster* than performing a traditional fixed-point iteration using a partitioned transition relation and early variable quantification. The technique of [17] does not have this property.

The compositional technique presented here shares ideas with partial model checking [1, 2, 16], but explicitly analyzes the structure of the system.

Finally, experiments by Anderson *et al.* [3] and Sreemani and Atlee [20] verified large software systems using SMV. The technique presented here significantly improves on the results we have obtained using SMV and makes it possible to verify larger systems.

Outline

The state/event model is described in section 2. Section 3 explains how the range of consistency checks performed by visualSTATE™ are reduced to two simple types of checks. Section 4 shows how state/event systems are encoded by ROB-DDs. The compositional technique and the dependency analysis is introduced in section 5, and further developed in section 6. The technique is evaluated in section 7 and section 8 draws some conclusions.

2 State/Event Systems

A state/event system consists of n machines $M_1, \ldots, M_n$ over an alphabet of input events E and an output alphabet O. Each machine M_i is a triple (S_i, s_i^0, T_i) of local states, an initial state, and a set of transitions. The set of transitions is a relation

$$T_i \subseteq S_i \times E \times G_i \times \mathcal{M}(O) \times S_i,$$

where $\mathcal{M}(O)$ is a multi-set of outputs, and G_i is the set of guards not containing references to machine i. These guards are generated from the following simple grammar for Boolean expressions:

$$g \quad ::= \quad l_j = p \mid \neg g \mid g \wedge g \mid tt.$$

The atomic predicate $l_j = p$ is read as "machine j is at local state p" and tt denotes a true guard. The global state set of the state/event system is the product of the local state sets: $S = S_1 \times S_2 \times \cdots \times S_n$. The guards are interpreted straightforwardly over S: for any $s \in S$, $s \models l_j = p$ holds exactly when the j'th component of s is p, i.e., $s_j = p$. The notation $g[s_j/l_j]$ denotes that s_j is substituted for l_j, with occurrences of atomic propositions of the form $s_j = p$ replaced by tt or $\neg tt$ depending on whether s_j is identical to p.

Considering a global state s, all guards in the transition relation can be evaluated. We define a version of the transition relation in which the guards have been evaluated. This relation is denoted $s \xrightarrow{e\ o}_i s_i'$ expressing that machine i when receiving input event e makes a transition from s_i to s_i' and generates output o. Formally,

$$s \xrightarrow{e\ o}_i s_i' \Leftrightarrow_{def} \exists g.\ (s_i, e, g, o, s_i') \in T_i \text{ and } s \models g.$$

Two machines can be combined into one. More generally if M_I and M_J are compositions of two disjoint sets of machines I and J, $I, J \subseteq \{1, \ldots, n\}$, we can combine them into one $M_{IJ} = M_I \times M_J = (S_{IJ}, (s_{IJ}^0), T_{IJ})$, where $S_{IJ} = S_I \times S_J$ and $s_{IJ}^0 = (s_I^0, s_J^0)$. The transition relation T_{IJ} is a subset of $S_{IJ} \times E \times G_{IJ} \times \mathcal{M}(O) \times S_{IJ}$, where G_{IJ} are the guards in the composite machine. By construction of T_{IJ}, the guards G_{IJ} will not contain any references to machines in $I \cup J$. To define T_{IJ}, we introduce the predicate $idle$:

$$idle(T_I, s_I, e) = \bigwedge \{\neg g \mid \exists o, s_I'.\ (s_I, e, g, o, s_I') \in T_I\},$$
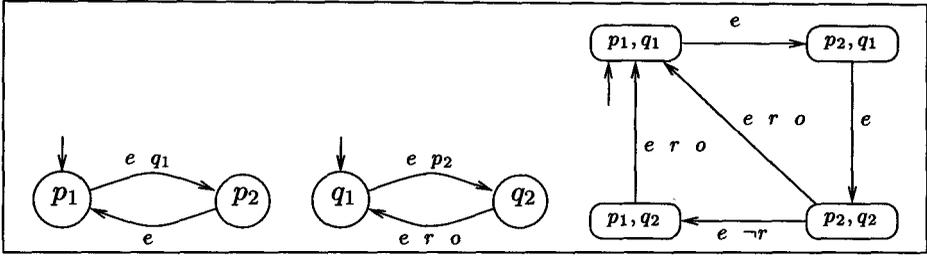
Fig. 1. Two state/event machines and the corresponding parallel combination. The guards, which formally should be of the form $l_j = p$, are simply written as the state p since the location l_j is derivable from the name of the state. The small arrows indicate the initial states. The reference to r is a requirement to a state in a third machine not shown.

which holds for states in which no transitions in M_I are enabled at state s_I when receiving the input event e. The transition relation T_{IJ} is defined by the following inference rules (the symbol $\uplus$ denotes multi-set union):

$$\frac{(s_I, e, g_1, o_1, s_I') \in T_I \quad (s_J, e, g_2, o_2, s_J') \in T_J}{((s_I, s_J), e, g, o_1 \uplus o_2, (s_I', s_J')) \in T_{IJ}} \quad g = g_1[s_J/l_J] \wedge g_2[s_I/l_I]$$

$$\frac{(s_I, e, g_1, o_1, s_I') \in T_I}{((s_I, s_J), e, g, o_1, (s_I', s_J)) \in T_{IJ}} \quad g = g_1[s_J/l_J] \wedge idle(T_J, s_J, e)[s_I/l_I]$$

$$\frac{(s_J, e, g_2, o_2, s_J') \in T_J}{((s_I, s_J), e, g, o_2, (s_I, s_J')) \in T_{IJ}} \quad g = idle(T_I, s_I, e)[s_J/l_J] \wedge g_2[s_I/l_I].$$

The rules show the synchronous behaviour of state/event systems. The first rule represents the case where there exists an enabled transition with input event e in both T_I and T_J and the resulting transition in T_{IJ} represents the synchronization on e. The other two cases occur if no enabled transition exists in either T_I or T_J. The state/event systems in figure 1 shows two machines and the parallel composition of them.

The full combination of all n machines yields a Mealy machine in which the transitions $s \xrightarrow{e\ o} s'$ are defined by

$$s \xrightarrow{e\ o} s' \iff \exists g. \ (s, e, g, o, s') \in \prod_{i=1}^{n} T_i \text{ and } g \text{ is true.}$$

3 Consistency Checks

The consistency checker in visualSTATE™ performs seven predefined types of checks, each of which can be reduced to verifying one of two types of properties. The first type is a *reachability property*. For instance, visualSTATE™ performs a

check for "conflicting transitions" i.e., it checks whether two or more transitions can become enabled in the same local state, leading to non-determinism. This can be reduced to questions of reachability by considering all pairs of guards g_1 and g_2 of transitions with the same local state s_i and input event e. A conflict can occur if a global state is reachable in which $(l_j = s_i) \wedge g_1 \wedge g_2$ is satisfied.

In total, five of the seven types of checks reduce to reachability checks. Four of these, such as check for transitions that are never enabled and check for states that are never reached, generate a number of reachability checks which is linear in the number of transitions, t. In the worst-case the check for conflicting transitions gives rise to a number of reachability checks which is quadratic in the number of transitions. However, in practice very few transitions have the same starting local state and input event, thus in practice the number of checks generated is much smaller than t.

The remaining two types of consistency checks reduce to a check for absence of *local deadlocks*. A local deadlock occurs if the system can reach a state in which one of the machines idles forever on all input events. This check is made for each of the n machines. In total at least $t + n$ checks have to be performed making the verification of state/event systems quite different from traditional model checking where typically only a few key properties are verified.

We attempt to reduce the number of reachability checks by performing an *implicational analysis* between the guards of the checks. If a guard g_1 implies another guard g_2 then clearly, if g_1 is reachable so is g_2. To use this information we start by sorting all the guards in ascending order of the size of their satisfying state space. In this way the most specific guards are checked first and for each new guard to be checked we compare it to all the already checked (and reachable) guards. If the new guard includes one of them, then we know that it is satisfiable. In our experiments, between 40% and 94% of the reachability checks are eliminated in this manner.

4 ROBDD Representation

This section describes how Reduced Ordered Binary Decision Diagrams (ROBDDs) [5] are used to represent sets of states and the transition relation. We also show how to perform a traditional forward iteration to construct the set of reachable states from which it is straightforward to check each of the reachability checks.

To construct the ROBDD $\widetilde{T}$ for the transition relation T, we first construct the local transition relations $\widetilde{T}_i$ for each machine M_i. The variables of the ROBDD represents an encoding of the input events, the current states, and the next-states. The variables are ordered as follows: The first $\|E\|$ variables encodes the input events E ($\|X\|$ denotes $\lceil \log_2 |X| \rceil$) and are denoted V_E. Then follows $2\|S_i\|$ variables $V_{i,1}, V'_{i,1}, \dots, V_{i,\|S_i\|}, V'_{i,\|S_i\|}$ encoding the current- (unprimed variables) and the next-states (primed variables) for machine i.

The transition relation $\widetilde{T}_i$ for machine i is constructed as an ROBDD predicate over these variables. The ROBDD for a transition $(s_i, e, g, o, s'_i) \in T_i$ is

constructed as the conjunction of the ROBDD encodings of s_i, e, g, and s_i'. (The outputs are not encoded as they have no influence on the reachable states of the system.) The encoding of s_i, e, and s_i' is straightforward and the encoding of the guard g is done by converting all atomic predicates $l_j = p$ to ROBDD predicates over the current-state variables for machine M_j and then performing the Boolean operations in the guard. The encoding of all transitions of machine i is obtained from the disjunction of the encoding of the individual transitions:

$$\widetilde{t}_i = \bigvee_{(s_i,e,g,o,s_i')\in T_i} \widetilde{s}_i \wedge \widetilde{e} \wedge \widetilde{g} \wedge \widetilde{s}_i',$$

where $\widetilde{e}$ is the ROBDD encoding of input event e and $\widetilde{s}_i$ and $\widetilde{s}_i'$ are the ROBDD encodings of the current-state s_i and next-state s_i', respectively.

To properly encode the global transition relation T, we need to deal with situations where no transitions of T_i are enabled. In those cases we want the machine i to stay in its current state. We construct an ROBDD neg_i representing that no transition is enabled by negating all guards in machine i (including the input events):

$$neg_i = \bigwedge_{(s_i,e,g,o,s_i')\in T_i} \neg(\widetilde{s}_i \wedge \widetilde{g} \wedge \widetilde{e}) .$$

The ROBDD equ_i encodes that machine i does not change state by requiring that the next-state is identical to the current-state:

$$equ_i = \bigwedge_{j=1}^{\|S_i\|} V_{i,j} \leftrightarrow V_{i,j}' .$$

The local transition relation for machine i is then:

$$\widetilde{T}_i = \widetilde{t}_i \vee (neg_i \wedge equ_i) .$$

The ROBDD $\widetilde{T}$ for the full transition relation is the conjunction of the local transition relations:

$$\widetilde{T} = \bigwedge_{i=1}^{n} \widetilde{T}_i .$$

One way to check whether a state s is reachable is to construct the reachable state space R. The construction of R can be done by a standard forward iteration of the transition relation, starting with the initial state s^0:

$$R_0 = \widetilde{s}^0$$
$$R_k = R_{k-1} \vee (\exists V, V_E. \; \widetilde{T} \wedge R_{k-1})[V/V']$$

where V is the set of current-state variables, V' is the set of next-state variables, and $(\cdots)[V/V']$ denotes the result of replacing all the primed variables in V' by their unprimed versions.

The construction of the full transition relation T can be avoided by using a *partitioned transition relation* [6] together with early variable quantification.

This is done by identifying sets I_j of transition relations that, when applied in the correct order, allows for early quantification of the state variables that no other transition relations depend on. If V_{I_j} are these variables and we have m sets, then we get:

$$R_0 = \widetilde{s}^0$$
$$R_k = R_{k-1} \vee (\exists V_E.\exists V_{I_1}. \bigwedge_{i \in I_1} \widetilde{T}_i \wedge \ldots \wedge (\exists V_{I_m}. \bigwedge_{i \in I_m} \widetilde{T}_i \wedge R_{k-1}))[V/V'].$$

Both approaches have been implemented and tested on our examples as shown in section 7. Here we see that the calculation of the reachable state space using the full transition relation is both fast and efficient for the small examples. However, for models with more than approximately 30 machines, both approaches fail to complete.

5 Compositional Backwards Reachability

The problems of forwards iteration can typically be solved by using a backwards reachability analysis. The verification task is to determine whether a guard g can be satisfied. Instead of computing the reachable state space and check that g is valid somewhere in this set, we start with the set of states in which g is valid and compute in a backwards iteration, states that can reach a state in which g is satisfied. The goal is to determine whether the initial state is among these states. Our novel idea is to perform the backwards iteration in a compositional manner considering only a minimal number of machines. Initially, only machines mentioned in g will be taken into account. Later also machines on which these depend will be included.

Notice that compared to the forwards iteration, this approach has an apparent drawback when performing a large number of reachability checks: instead of just *one* fixed-point iteration to construct the reachable state space R (and then trivially verify the properties), a new fixed-point iteration is necessary for *each* property that is checked. However, our experiments clearly demonstrate that when using a compositional backwards iteration, each of the fixed-point iterations can be completed even for very large models whereas the forwards iteration fails to complete the construction of R for even medium sized models.

To formalize the backwards compositional technique, we need a semantic version of the concept of dependency. For a subset of the machines $I \subseteq \{1, \ldots, n\}$, two states $s, s' \in S$ are *I-equivalent*, written $s =_I s'$, if for all $i \in I$, $s_i = s'_i$ (the primes are here used to denote another state and is not related to the next-states). If a subset P of the reachable states S only is constrained by components in some index set I we can think of P as having I as a *sort*. This leads to the following definition: a subset P of S is *I-sorted* if for all $s, s' \in S$,

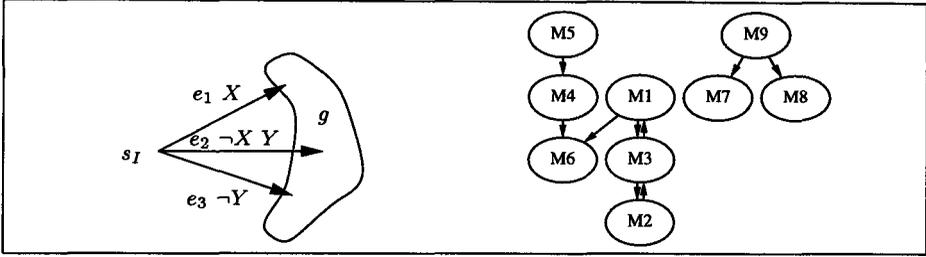$$s \in P \text{ and } s =_I s' \Rightarrow s' \in P.$$

Fig. 2. The left figure is an example showing the effect of $B_I(g)$. If X is the guard $l_j = p$ and Y the guard $l_k = q$ with $j, k \notin I$ then the transitions from s_I seem to depend on machines M_j and M_k outside I. However, the guards X, $\neg X\ Y$, and $\neg Y$ together span all possibilities and therefore, by selecting either e_1, e_2, or e_3, the state s_I can reach g irrespective of the states of the machines M_j and M_k. The right figure illustrates the dependencies between 9 state machines taken from a real example (the example "HI-FI" of section 7). An arrow from one machine M_i to another M_j indicates the existence of a transition in M_i with a guard that depends on a state in machine M_j.

As an example, consider a guard g which mentions only machines 1 and 3. The set of states defined by g is I-sorted for any I containing 1 and 3.[1] Another understanding of the definition is that a set P is I-sorted if it is independent of the machines in the complement $\overline{I} = \{1, \ldots, n\} \setminus I$.

From an I-sorted set g we perform a backwards reachability computation by including states which, irrespective of the states of the machines in $\overline{I}$, can reach g. One backward step is given by the function $B_I(g)$ defined by:

$$B_I(g) = \{s \in S \mid \forall s' \in S.\ s =_I s' \Rightarrow \tag{1}$$
$$\exists e, o, s''.\ s' \xrightarrow{e\ o} s'' \text{ and } s'' \in g\}.$$

By definition $B_I(g)$ is I-sorted. The set $B_I(g)$ is the set of states which independently of machines in $\overline{I}$, by some input event e, can reach a state in g. Observe that $B_I(g)$ is monotonic in both g and I. Figure 2 shows how a state s_I of a machine is included in $B_I(g)$ although it syntactically seems to depend on machines outside I.

By iterating the application of B_I, we can compute the minimum set of states containing g and closed under application of B_I. This is the minimum fixed-point $\mu X.g \cup B_I(X)$, which we refer to as $B_I^*(g)$. Note that $B_{\{1,\ldots,n\}}^*(g)$ becomes the desired set of states which can reach g.

A set of indices I is said to be *dependency closed* if none of the machines in I depend on machines outside I. Formally, I is *dependency closed* if for all $i \in I$, states s', s, s_i, input events e, and outputs o, $s \xrightarrow{e\ o}_i s_i$ and $s' =_I s$ implies $s' \xrightarrow{e\ o}_i s_i$.

---

[1] If the guard is self-contradictory (always false), it will be I-sorted for any I. This reflects the fact that the semantic sortedness is more precise than syntactic occurrence.

The basic properties of the sets $B_I^*(g)$ are captured by the following lemma:

Lemma 1 (Compositional Reachability Lemma). *Assume g is an I-sorted subset of S. For all subsets of machines I, J with $I \subseteq J$ the following holds:*

$(i).\ B_I^*(g) \subseteq B_J^*(g)$

$(ii).\ B_J^*(g) = B_J^*(B_I^*(g))$

$(iii).\ I\ \text{dependency closed} \;\Rightarrow\; B_I^*(g) = B_J^*(g).$

The results of the lemma are applied in the following manner. To check whether a guard g is reachable, we first consider the set of machines I_1 syntactically mentioned in g. Clearly, g is I_1-sorted. We then compute $B_{I_1}^*(g)$. If the initial state s^0 belongs to $B_{I_1}^*(g)$, then by (i) $s^0 \in B_{\{1,\ldots,n\}}^*(g)$ and therefore g is reachable from s^0 and we are done. If not, we extend I_1 to a larger set of machines I_2 by adding machines that are syntactically mentioned on guards in transitions of machines in I_1. We then reuse $B_{I_1}^*(g)$ to compute $B_{I_2}^*(g)$ as $B_{I_2}^*(B_{I_1}^*(g))$ which is correct by (ii). We continue like this until s^0 has been found in one of the sets or an index set I_k is dependency closed. In the latter case we have by (iii) $B_{I_k}^*(g) = B_{\{1,\ldots,n\}}^*(g)$ and g is unreachable unless $s^0 \in B_{I_k}^*(g)$.

As an example, assume that we want to determine whether the guard $g = (l_1 = p \wedge l_3 \neq q)$ is reachable in the example of figure 2 (right). The initial index set is $I_1 = \{1,3\}$. If this is not enough to show g reachable, the second index set $I_2 = \{1,3,7,2\}$ is used. Since this set is dependency closed, g is reachable if and only if the initial state belongs to $B_{I_2}^*(B_{I_1}^*(g))$.

The above construction is based on a backwards iteration. A dual version of B_I for a forwards iteration could be defined. However, such a definition would not make use of the dependency information since s^0 is only I-sorted for $I = \{1,\ldots,n\}$. Therefore all machines would be considered in the first fixed-point iteration reducing it to the complete forwards iteration mentioned in the previous section.

Seemingly, the definition of $B_I(g)$ requires knowledge of the global transition relation and therefore does not seem to yield any computational advantage. However, as explained below, using ROBDDs this can be avoided leading to an efficient computation of $B_I(g)$. The ROBDD $\widetilde{B}_I(\widetilde{g})$ representing one iteration backwards from the states represented by the ROBDD $\widetilde{g}$ can be constructed immediately from the definition (1):

$$\widetilde{B}_I(\widetilde{g}) = \forall V_{\bar{I}}.\ \exists V_E, V'.\ \widetilde{T}\ \wedge\ \widetilde{g}[V'/V] \tag{2}$$

where $\widetilde{g}[V'/V]$ is equal to $\widetilde{g}$ with all variables in V replaced by their primed versions. It is essential to avoid building the global transition relation $\widetilde{T}$. This is done by writing $\exists V'$ as $\exists V_I'.\exists V_{\bar{I}}'$ and $\widetilde{T} = \widetilde{T}_I \wedge \widetilde{T}_{\bar{I}}$ where $\widetilde{T}_I = \bigwedge_{i \in I} \widetilde{T}_i$. This allows us to push the existential quantification of $V_{\bar{I}}'$ to $\widetilde{T}_{\bar{I}}$ since g is I-sorted and thus independent of the variables in $V_{\bar{I}}'$. As $\exists V_{\bar{I}}'.\widetilde{T}_{\bar{I}}$ is a tautology, equation (2) reduces to:

$$\widetilde{B}_I(\widetilde{g}) = \forall V_{\bar{I}}.\ \exists V_E, V_I'.\ \widetilde{T}_I\ \wedge\ \widetilde{g}[V'/V]$$

which only uses the local transition relations for machines in I. Each T_i refers only to primed variables in V_i', allowing early variable quantification for each machine individually:

$$\widetilde{B}_I(\widetilde{g}) = \forall V_{\bar{I}}. \ \exists V_E. \ \exists V_1'.\widetilde{T}_1 \wedge (\exists V_2'.\widetilde{T}_2 \wedge \cdots \wedge (\exists V_k'.\widetilde{T}_k \wedge \widetilde{g}[V'/V]) \cdots).$$

This equation efficiently computes one step in the fixed-point iteration constructing $\widetilde{B}_I^*(\widetilde{g})$.

Notice, that the existential quantifications can be performed in any order. We have chosen the order in which the machines occur in the input, but other orders may exist which improves performance.

6   Local Deadlock Detection

In checking for local deadlocks we use a construction similar to backwards reachability. To make the compositional backwards lemma applicable we work with the notion of a machine being *live* which is the exact dual of having a local deadlock. In words, a machine is live if it always is the case that there exists a way to make the machine move to a new local state. Formally, a global state s is *live* for machine i if there exists a sequence of states $s^1, s^2, \ldots, s^k$ with $s = s^1$ and $s^j \xrightarrow{e\ o} s^{j+1}$ (for some e and o) such that $s_i^k \neq s_i^1$. Machine i is live if all reachable states are live. A simple example of a state/event system with a local deadlock is shown in figure 3.
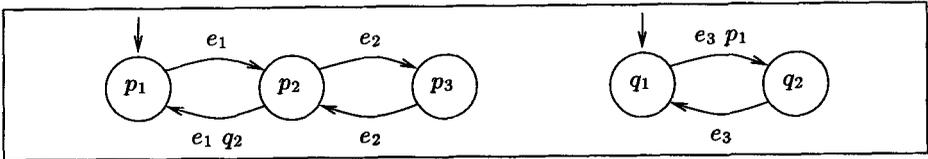

Fig. 3. A state/event system with a local deadlock. The global state $s = (p_2, q_1)$ is not live for the machine to the right since for all input events the guard p_1 remains false. The state s is reachable (e.g., by initially receiving e_1) and thus the machine to the right has a local deadlock.

The check is divided into two parts. First, the set of all live states L_i^* for machine i is computed. Second, we check that all reachable states are in L_i^*. A straightforward but inefficient approach would be to compute the two sets and check for inclusion. However, we will take advantage of the compositional construction used in the backwards reachability in both parts of the check.

Similar to the definition of $B_I(g)$, we define $L_{I,i}(X)$ to be the set of states that are immediately live for machine $i \in I$ independently of the machines outside I or leads to states in X (which are states assumed to be live for machine i):

$$L_{I,i}(X) = \{ s \in S \mid \forall s'.s =_I s' \ \Rightarrow \qquad\qquad (3)$$
$$\exists e, o, s''. \ s' \xrightarrow{e\ o} s'' \text{ and } (s_i \neq s_i'' \text{ or } s'' \in X) \}.$$

Notice, that compared to definition (1) the only difference is the extra possibility that the state is immediately live, i.e., $s_i \neq s_i''$. The set of states that are live for machine i independently of machines outside I is then the set $L_{I,i}^*(\emptyset)$ where $L_{I,i}^*(Y)$ is the minimum fixed point defined by $L_{I,i}^*(Y) = \mu X.Y \cup L_{I,i}(X)$.

The three properties of lemma 1 also holds for $L_{I,i}^*(Y)$ when Y is I-sorted. If I is dependency closed it follows from property (iii) of the lemma that $L_{I,i}^*(\emptyset)$ equals $L_{\{1,\ldots,n\},i}^*(\emptyset)$ which is precisely the set of live states of machine i. This gives an efficient way to compute the sets $L_{I,i}^*(\emptyset)$ for different choices of I. We start with I_1 equal to $\{i\}$ and continue with larger I_k's exactly as for the backwards reachability. The only difference is the termination conditions. One possible termination case is if $L_{I_k,i}^*(\emptyset)$ becomes equal to S for some k. In that case it is trivial that the set of reachable states is contained in $L_{I_k,i}^*(\emptyset)$. From the monotonicity property (i) of the lemma it follows that machine i is live and thus free of local deadlocks. The other termination case is when I_k becomes dependency closed. Then we have to check whether there exists reachable states not in $L_{I_k,i}^*(\emptyset)$. This is done by a compositional backwards reachability check with $g = S \setminus L_{I_k,i}^*(\emptyset)$.

7 Experimental Results

The technique presented above has been used on a range of real industrial state/event systems and a set of systems constructed by students in a course on embedded systems. The examples are all constructed using visualSTATE™ [18]. They cover a large range of different applications and are structurally highly irregular.

The examples HI-FI, AVS, FLOW, MOTOR, INTERVM, DKVM, MULTICD, TRAIN1 and TRAIN2 are all industrial examples. HI-FI is the control part of an advanced compact hi-fi system, AVS is the control part of an audio-video system, FLOW is the control part of a flow meter, MOTOR is a motor control, INTERVM and DKVM are advanced vending machines, MULTICD is the control of a multi-volume CD player, and TRAIN1 and TRAIN2 are both independent subsystems of a train simulator. The remaining examples are constructed by students. The VCR is a simulation of a video recorder, CYBER is an alarm clock, JVC is the control of a compact hi-fi system, VIDEO is a video player, and VOLVO is a simulation of the complete functionality of the dashboard of a car. The characteristics of the state/event systems are shown in table 1.

The experiments were carried out on a 166 MHz Pentium PC with 32 MB RAM running Linux. To implement the ROBDD operations, we have constructed our own ROBDD package which is comparable to state-of-the-art packages in terms of performance. In all experiments we limit the total number of ROBDD nodes to one million corresponding to 20 MB of memory. We check for each transition whether the guard is reachable and whether it is conflicting with other transitions. Furthermore, we check for each machine whether it has a local deadlock. The total runtime and memory consumption for these checks are shown in table 2. The total number of checks is far from the quadratic worst-case, which

Table 1. The state/event systems used in the experiments. The last two columns show the size of the declared and reachable state space. The size of the declared state space is the product of the number of local states of each machine. The reachable state space is only known for those systems where the forwards analysis completes.

System	Machines	Local states	Transitions	Declared	Reachable
INTERVM	6	182	745	10^6	15144
VCR	7	46	85	10^5	1279
CYBER	8	19	98	10^3	240
JVC	8	25	106	10^4	352
DKVM	9	55	215	10^6	377568
HI-FI	9	59	373	10^7	1416384
FLOW	10	232	1146	10^5	17040
MOTOR	12	41	295	10^6	34560
AVS	12	66	1737	10^7	1438416
VIDEO	13	74	268	10^8	1219440
VOLVO	20	84	196	10^{11}	$9.2 \cdot 10^9$
MULTICD	28	145	3128	10^{17}	$3.8 \cdot 10^{16}$
TRAIN1	373	931	2988	10^{136}	−
TRAIN2	1421	3204	11166	10^{476}	−

Table 2. The runtime and memory consumptions of the experiments. The second column of the table shows the total number of guards that are checked for reachability after this number has been reduced by the implicational analysis. The forward columns show results using a forward iteration with a full and a partitioned transition relation. The backward columns show the results of a backwards iteration using the full transition relation and the compositional backwards reachability. The visualSTATE column shows the runtimes obtained using an explicit state enumeration as implemented in version 3.0 of visualSTATE™. A "−" denotes that we ran out of memory or the runtime exceeded two hours without finishing.

| System	Guards checked	Forward Full Sec	MB	Partitioned Sec	MB	Backward Full Sec	MB	Compositional Sec	MB	visualSTATE Sec
INTERVM	185	1.7	6	6.1	6	6.5	6	22.5	6	22
VCR	50	0.3	6	0.6	5	0.9	6	0.7	7	< 1
CYBER	16	0.2	6	0.3	5	0.2	6	0.2	6	< 1
JVC	22	0.2	6	0.3	5	0.3	6	0.2	6	< 1
DKVM	63	0.8	6	16.0	6	1.7	6	3.4	6	409
HI-FI	120	1.2	6	25.2	6	6.2	6	7.2	6	1200
FLOW	230	1.2	6	3.7	6	8.8	6	10.8	6	22
MOTOR	123	1.0	6	11.6	6	4.9	6	3.2	6	32
AVS	173	8.9	6	126.4	8	9.7	6	13.0	6	3780
VIDEO	122	1.3	6	38.7	6	7.4	6	2.6	6	−
VOLVO	83	4.5	9	1271.3	5	14.3	6	2.6	6	−
MULTICD	199	4.9	6	32.4	5	11.0	6	4.2	6	−
TRAIN1	1335	−	−	−	−	−	−	38.3	9	−
TRAIN2	4708	−	−	−	−	−	−	2918.7	45	−
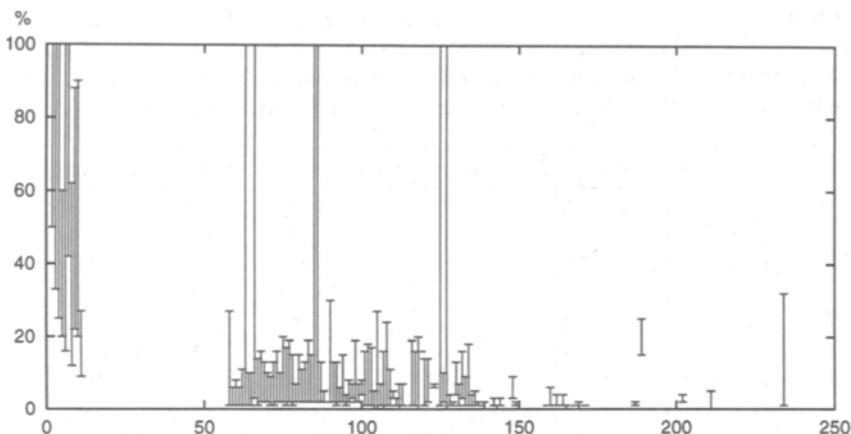
Fig. 4. The fraction of machines actually used in the compositional backwards reachability analysis of the guards of the largest system TRAIN2. For each size of dependency closed set, a line between is drawn between the minimum and maximum fraction of machines used in verifying guards with dependency closed sets of that size. For instance, for the guards with dependency closed sets with 234 machines (the right-most line) only between 1% and 32% of the machines are needed to prove that the guard is reachable.

supports the claim that in practice only very few checks are needed to check for conflicting rules (see section 3).

As expected, the forwards iteration with full transition relation is efficient for smaller systems. It is remarkable that the ROBDD technique is superior to explicit state enumeration even for systems with a very small number of reachable states. Using the partitioned transition relation in the forwards iteration works poorly.

For the two largest systems, only the compositional backwards technique succeeds. In fact, for the four largest systems it is the most efficient and for the small examples it has performance comparable to the full forward technique. This is despite the fact that the number of checks is high and the backward iterations must be repeated for each check. From the experiments it seems that the compositional backwards technique is better than full forwards from somewhere around 20 machines.

In order to understand why the compositional backwards technique is successful we have analyzed the largest system TRAIN2 in more detail, see figure 4. For each guard we have computed the size of its smallest enclosing dependency closed set of machines. During the backwards iterations we have kept track of how many times the set of machines I (used in $B_I^*(g)$) need to be enlarged and how many machines was contained in the set I when the iteration terminated. The dependency closed sets of cardinality 63, 66, 85, 86, 125, 127 all contain at least one machine with a guard that is unreachable. As is clearly seen from the figure, in these cases the iteration has to include the entire dependency closed set in order to prove that the initial state cannot reach the guard. But even

then much is saved, as no more than 234 machines out of a possible 1421 is ever included. In fact, *only* in the case of unreachable guards is more than 32% of the machines in a dependency closed set ever needed (ignoring the small dependency closed sets with less than 12 machines). A reduction of 32% amounts to a reduction in runtime much larger than a third due to the potential exponential growth of the ROBDD representation in the number of transition relations $\widetilde{T_i}$.

8   Conclusion

We have presented a verification problem for state/event systems which is characterized by a large number of reachability checks. A compositional technique has been presented which significantly improves on the performance of symbolic model checking for state/event systems. This has been demonstrated on a variety of industrial systems several of which could not be verified using traditional symbolic model checking, e.g., using SMV.

We expect that these results can be translated to other models of embedded control systems than the state/event model, as for example StateCharts [15]. Furthermore, the two types of checks are of a general nature. Clearly, reachability is a key property which captures a number of important properties of a system. Moreover, the check for local deadlock shows how properties requiring nesting of fixed points can be checked efficiently with the compositional backwards analysis. Thus, it seems straightforward to implement more general checks as expressed in, for instance, CTL [9].

References

1. H. R. Andersen, J. Staunstrup, and N. Maretti. Partial model checking with ROBDDs. In E. Brinksma, editor, *Proceedings of TACAS'97*, volume 1217 of *LNCS*, pages 35–49. Springer-Verlag, April 1997.
2. H.R. Andersen. Partial model checking (extended abstract). In *Proceedings, Tenth Annual IEEE Symposium on Logic in Computer Science*, pages 398–407, La Jolla, San Diego, 26–29 July 1995. IEEE Computer Society Press.
3. R.J. Anderson, P. Beame, S. M. Burns, W. Chan, F. Modugno, D. Notkin, and J.D. Reese. Model checking large software specifications. In D. Garlan, editor, *SIGSOFT '96. Proceedings of the Fourth ACM SIGSOFT Symposium on the Foundations of Software Engineering*, pages 156–66, San Francisco, 1996. ACM.
4. F. Balarin and A.L. Sangiovanni-Vincentelli. An iterative approach to language containment. In C. Courcoubetis, editor, *CAV'93. 5th International Conference on Computer Aided Verification*, volume 697 of *LNCS*, pages 29–40, Berlin, 1993. Springer-Verlag.
5. R.E. Bryant. Graph-based algorithms for Boolean function manipulation. *IEEE Transactions on Computers*, 8(C-35):677–691, 1986.
6. J. R. Burch, E. M. Clarke, and D. E. Long. Symbolic model checking with partitioned transition relations. In A. Halaas and P. B. Denyer, editors, *Proc. 1991 Int. Conf. on VLSI*, August 1991.

7. J. R. Burch, E. M. Clarke, K. L. McMillan, D. L. Dill, and L. J. Hwang. Symbolic model checking: 10^{20} states and beyond. In *Proceedings, Fifth Annual IEEE Symposium on Logic in Computer Science*, pages 428–439. IEEE Computer Society Press, 1990.

8. J.R. Burch, E.M. Clarke, D.E. Long, K.L. McMillan, and D.L. Dill. Symbolic model checking for sequential circuit verification. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 13(4):401–424, 1994.

9. E.M. Clarke, E.A. Emerson, and A.P. Sistla. Automatic verification of finite-state concurrent systems using temporal logic specifications. *ACM Transactions on Programming Languages and Systems*, 8(2):244–263, 1986.

10. E.M. Clarke, O. Grumberg, and D.E. Long. Model checking and abstraction. *ACM Transactions on Programming Languages and Systems*, 1994.

11. E.M. Clarke, D.E. Long, and K.L. McMillan. Compositional model checking. In *Proceedings, Fourth Annual Symposium on Logic in Computer Science*, pages 353–362, Asilomar Conference Center, Pacific Grove, California, June 5–8 1989. IEEE Computer Society Press.

12. O. Coudert, C. Berthet, and J. C. Madre. Verification of synchronous sequential machines based on symbolic execution. In J. Sifakis, editor, *Automatic Verification Methods for Finite State Systems. Proceedings*, volume 407 of *LNCS*, pages 365–373. Springer-Verlag, 1989.

13. O. Coudert, J. C. Madre, and C. Berthet. Verifying temporal properties of sequential machines without building their state diagrams. In E.M. Clarke and R.P. Kurshan, editors, *CAV'90. Workshop on Computer-Aided Verification.*, pages 75–84, Rutgers, New Jersey, 1990. American Mathematical Society.

14. D. Geist and I. Beer. Efficient model checking by automated ordering of transition relation partitions. In D.L. Dill, editor, *CAV'94. 6th International Conference on Computer Aided Verification*, volume 818 of *LNCS*, pages 299–310, Stanford, 1994. Springer-Verlag.

15. David Harel. STATECHARTS: A visual formalism for complex systems. *Science of Computer Programming*, 8(3):231–274, June 1987.

16. K. J. Kristoffersen, F. Laroussinie, K. G. Larsen, P. Patterson, and W. Yi. A compositional proof of a read-time mutual exclusion protocol. In M. Bidoit and M. Dauchet, editors, *Proceedings of TAPSOFT '97: Theory and Practice of Software Development*, volume 1214 of *LNCS*, pages 565–579. Springer-Verlag, 1997.

17. W. Lee, A. Pardo, J.-Y. Jang, G. Hachtel, and F. Somenzi. Tearing based automatic abstraction for CTL model checking. In *1996 IEEE/ACM International Conference on Computer-Aided Design*, pages 76–81, San Jose, CA, 1996. IEEE Comput. Soc. Press.

18. Beologic® A/S. *visualSTATE™ 3.0 User's Guide*, 1996.

19. K.L. McMillan. *Symbolic Model Checking*. Kluwer Academic Publishers, Norwell Massachusetts, 1993.

20. T. Sreemani and J.M. Atlee. Feasibility of model checking software requirements: a case study. In *COMPASS '96. Proceedings of the Eleventh Annual Conference on Computer Assurance*, pages 77–88, New York, USA, 1996. IEEE.