Modular Analysis of Tree-Topology Models *

Laure Petrucci¹ and Michał Knapik²

¹ LIPN, CNRS UMR 7030, Université Sorbonne Paris Nord, 99 av. J-B. Clément, 93430 Villetaneuse, France laure.petrucci@lipn.univ-paris13.fr
² Institute of Computer Science, Polish Academy of Sciences, Jana Kazimierza 5, 01-248 Warsaw, Poland m.knapik@ipipan.waw.pl

Abstract. We investigate networks of automata that synchronise over common action labels. A graph synchronisation topology between the automata is defined in such a way that two automata are connected iff they can synchronise over an action. We show a very effective reduction of networks of automata with tree-like synchronisation topologies. The reduction preserves a certain form of reachability, but not safety. The procedure is implemented in an open-source tool.

1 Introduction

Networks of various flavours of finite automata are the usual choice of formalism when modeling complex systems such as protocols. This approach also plays well with the divide-and-conquer paradigm, as the investigated system can be divided into components modeled with various degree of granularity. However, the cost of computing the synchronised product of these submodules can be prohibitive: in practice the size of the state-space grows exponentially with the number of components.

In this paper we tackle the problem of computing of a part of the state-space of the entire synchronised product in such a way that a certain version of reachability is preserved. At this stage we only deal with systems that exhibit tree-like synchronisation structure and consist of live-reset automata. Namely, each component can synchronise via shared upstream actions with a single other module (its parent) after which it resets, i.e. returns to the initial state. We propose a bottom-up reduction based on the observation that any execution of the entire system can be rewritten in a reachability-preserving way into a sequence of interactions between components and their parents followed by upstream synchronisations. Thus, the reduced model is constructed by creating synchronised products of pairs consisting of a component and its parent. The size of the state-space of the resulting automaton is much smaller than the product of the entire network.

The theory has been implemented in an open-source tool [1].

*This work was supported by the PICS CNRS/PAN project PARTIES and by PAN. Michał Knapik is supported by POLLUX VoteVerif project.
2 Tree Synchronisation Systems

In this section we recall the basic notions of networks of Labelled Transition Systems and their synchronisation topologies. We also introduce and explain the restrictions on the models assumed in this paper. In what follows let $\mathcal{PV}$ denote the set of propositions.

Definition 1 (Labelled Transition System). A Labelled Transition System (LTS) is a tuple $M = \langle S, s^0, Acts, \rightarrow, \mathcal{L} \rangle$ where:
1. $S$ is a finite set of states and $s^0 \in S$ the initial state;
2. Acts is a finite set of action names;
3. $\rightarrow \subseteq S \times Acts \times S$ is a transition relation;
4. $\mathcal{L} : S \rightarrow 2^{\mathcal{PV}}$ assigns to each state a set of propositions that hold therein.

We usually write $s \xrightarrow{act} s'$ instead of $(s, act, s') \in \rightarrow$. We also denote $\mathit{acts}(M) = Acts$ and $\mathit{states}(M) = S$. A run in LTS $M$ is an infinite sequence of states and actions $\rho = s^0act^0s^1act^1\ldots$ s.t. $s^i \xrightarrow{act^i} s^{i+1}$ for all $i \geq 0$. By $\mathit{Runs}(M, s)$ we denote the set of all the runs starting from state $s \in S$; if $s$ is the initial state, we simply write $\mathit{Runs}(M)$.

2.1 LTS Nets and Synchronisation Topologies

Both the commercial and research model checkers such as SPIN, UPPAAL or IMITATOR \cite{spin,uppaal,imitator} typically expect the systems described in a form of interacting modules. Concurrent transitions via common actions (or channels) are one of the most basic synchronisation primitives \cite{sync}.

Definition 2 (Asynchronous Product). Let $M_i = \langle S_i, s^0_i, \rightarrow_i, Acts_i, \mathcal{L}_i \rangle$ be LTS, for $i \in \{1, 2\}$. The asynchronous product of $M_1$ and $M_2$ is the LTS $M_1 || M_2 = \langle S_1 \times S_2, (s^0_1, s^0_2), \rightarrow, Acts_1 \cup Acts_2, \mathcal{L}_1 \cup \mathcal{L}_2 \rangle$ with the transition rule defined in the usual way:

\[
\begin{align*}
act & \in Acts_1 \setminus Acts_2 \land s_1 \xrightarrow{act_1} s'_1 & (s_1, s_2) \xrightarrow{act_1} (s'_1, s_2) \\
act & \in Acts_2 \setminus Acts_1 \land s_2 \xrightarrow{act_2} s'_2 & (s_1, s_2) \xrightarrow{act_2} (s_1, s'_2) \\
act & \in Acts_1 \cap Acts_2 \land s_1 \xrightarrow{act_1} s'_1 \land s_2 \xrightarrow{act_2} s'_2 & (s_1, s_2) \xrightarrow{act} (s'_1, s'_2)
\end{align*}
\]

The above definition is naturally extended to an arbitrary number of components, where we sometimes write $\prod_{i=0}^n M_i$ instead of $M_1 || \ldots || M_n$.

The synchronisation topology is an undirected graph that records how components synchronise with one another.
Definition 3 (Synchronisation Topology). A synchronisation topology \((ST)\) is a tuple \(G = (\text{Net}, T)\), where \(\text{Net} = \{\mathcal{M}_i\}_{i=1}^n\) is a set of LTSs for \(i \in \{1, \ldots, n\}\), and \(T \subseteq \text{Net} \times \text{Net}\) is s.t. \((\mathcal{M}_i, \mathcal{M}_j) \in T\) iff \(i \neq j\) and \(\text{Acts}_i \cap \text{Acts}_j \neq \emptyset\).

Note that \(T\) is induced by \(\text{Net}\). Thus, with a slight notational abuse we sometimes treat \(G\) as \(\text{Net}\). Moreover, we put \(\text{acts}(G) = \bigcup_{i=1}^n \text{acts}(\mathcal{M}_i)\).

In what follows we assume that \(G\) is a tree with the root \(\text{root}(G)\). Moreover, for each \(\mathcal{M} \in \text{Net}\) by \(\text{parent}(\mathcal{M})\) we denote its parent (we assume \(\text{parent}(\text{root}(G)) = \emptyset\)) and by \(\text{children}(\mathcal{M})\) we mean the set of its children. By \(\text{upacts}(\mathcal{M})\) (resp., \(\text{downacts}(\mathcal{M})\)) we denote the set of actions via which \(\mathcal{M}\) synchronises with its parent (children, resp.). For each \(act \in \text{downacts}(\mathcal{M})\) by \(\text{snd}(\mathcal{M}, act)\) we denote the component \(\mathcal{M}' \in \text{children}(\mathcal{M})\) s.t. \(act \in \text{upacts}(\mathcal{M}')\). Thus, \(\text{snd}(\mathcal{M}, act)\) is the child of \(\mathcal{M}\) that synchronises with \(\mathcal{M}\) over \(act\). If \(\mathcal{M}\) is clear from the context, we simply write \(\text{snd}(act)\).

The local, unsynchronised actions of \(\mathcal{M}\) are defined as \(\text{locacts}(\mathcal{M}) = \text{acts}(\mathcal{M}) \setminus (\text{downacts}(\mathcal{M}) \cup \text{upacts}(\mathcal{M}))\). For brevity, whenever we refer to a state or transition of \(G\) we mean a state or transition of \(||_{i=0}^n \mathcal{M}_i\). We also extend the notion of runs to synchronisation topologies: \(\text{Runs}(G, s) = \text{Runs}(||_{i=0}^n \mathcal{M}_i, s)\) for each \(s \in \text{states}(||_{i=0}^n \mathcal{M}_i)\).

We are interested in networks whose all components share a similar, simple structure. Namely, we say that an \(\text{LTS, M}\) is live-reset if every run \(\rho \in \text{Runs}(\mathcal{M})\) is s.t. executing any action from \(\text{upacts}(\mathcal{M})\) leads to the initial state. Intuitively, \(\mathcal{M}\) can freely synchronise with its children and execute local actions but resets once synchronising with the parent. If every component of an \(ST\) \(G\) is live-reset then we say that \(G\) is live-reset.

Example 1. Figure 1 presents a small tree \(ST\) \(G_x\) with the root \(R\) and two children \(\mathcal{M}_1\) and \(\mathcal{M}_2\). The auxiliary symbols \(?!\) are syntactic sugar, used to distinguish between \(\text{upacts}\) and \(\text{downacts}\). Here, \(\text{upacts}(R) = \emptyset, \text{downacts}(R) = \{\text{open}, \text{chooseL}, \text{chooseR}\}\), and \(\text{locacts}(R) = \{\text{beep}\}\). Similarly, \(\text{upacts}(\mathcal{M}_1) = \{\text{open}\}, \text{upacts}(\mathcal{M}_2) = \{\text{chooseL}, \text{chooseR}\}\), \(\text{downacts}(\mathcal{M}_1) = \text{locacts}(\mathcal{M}_1) = \{\text{beep}\}\), \(\text{downacts}(\mathcal{M}_2) = \{\text{beep}\}\).
downacts($M_2$) = $\emptyset$, and locacts($M_2$) = $\tau$. All the components of the model are live-reset.

Let $G = \langle Net, T \rangle$ be a ST. For each $M \in Net$ by $G_M$ we denote the ST induced by the subtree of $G$ rooted in $M$. Let $M \subseteq Net$ and $\rho^*$ be a prefix of some run $\in \text{Runs}(G)$ s.t. $\rho^* = s^0act^0s^1act^1 \ldots$. By $\rho^* \downarrow (M)$ we denote the projection of $\rho^*$ to the product of components in $M$, i.e., the result of transforming $\rho^*$ by (1) firstly projecting each $s^i$ on the components in $M$; (2) secondly, removing the actions that do not belong to $M$, together with their sources.

**Example 2.** Consider a sequence:

$$\eta = (r_0,s_0,t_0)\tau(r_0,s_0,t_1)\tau(r_0,s_0,t_2)\text{open}(r_1,s_0,t_2)$$

chooseR(r_4,s_0,t_0)\tau\text{chooseL}(r_0,s_0,t_0).$$

Here, we have $\rho^* \downarrow (R, M_1) = (r_0,s_0)\text{open}(r_1,s_0)\text{chooseR}(r_4,s_0)\text{chooseL}(r_0,s_0)$.

### 3 Reducing Live-Reset Trees

In this section we show how to create for a given synchronisation topology $G$ of live-reset components an LTS that preserves reachability. The procedure is presented in two steps. Firstly, we show how to build an LTS for two-level trees. Secondly, we show how to modify the former to deal with trees of arbitrary height in a bottom-up manner.

#### 3.1 Reduction for Two-level Trees

Throughout this subsection let $G$ be a live-reset tree ST with components $Net = \{R, M_1, \ldots, M_n\}$ s.t. root($G$) = $R$ and children($R$) = $\{M_1, \ldots, M_n\}$. Moreover, let $R = \langle S_R, s^0_R, \text{Acts}_R, \rightarrow_R, L_R \rangle$ and $M_i = \langle S_i, s^0_i, \text{Acts}_i, \rightarrow_i, L_i \rangle$, for $i \in \{1, \ldots, n\}$. We employ the observations on the nature of synchronisations with live-reset components in the following definition.

**Definition 4 (Unreduced Sum-of-squares Product).** Let $SQ^u(G) = \langle S^u_{sq}, s^0_{sq}, \text{Acts}_{sq}, \rightarrow_{sq}, L_{sq} \rangle$ be an LTS s.t.:

- $S^u_{sq} = \bigcup_{i=1}^n M_i \times R.$
- $s^0_{sq} \notin S^u_{sq}$ is a fresh initial state.
- $\text{Acts}_{sq} = \text{acts}(G) \cup \{\epsilon\}$, where $\epsilon \notin \text{acts}(G)$ is a fresh, silent action.
- The transition relation $\rightarrow_{sq}$ is defined as follows:

  - $s^0_{sq} \xrightarrow{\epsilon}_{sq} (s^0_i, s^0_R), \text{for all } i \in \{1, \ldots, n\}$; intuitively, using the new initial state of $SQ(G)$ and $\epsilon$-transitions we can visit the initial state of any square product $M_i \times R$. 

4
• If $s_i \xrightarrow{\text{act}_{i}} s_i'$ and $\text{act} \in \text{locacts}(M_i)$, then $(s_i, s_R) \xrightarrow{\text{act}_q} (s_i', s_R)$, for each $s_R \in S_R$; similarly, if $s_{R} \xrightarrow{\text{act}_{R}} s_{R}'$ and $\text{act} \in \text{locacts}(R)$, then $(s_i, s_{R}) \xrightarrow{\text{act}_q} (s_i, s_{R}')$, for each $s_i \in S_i$. Thus, the square products are fully asynchronous over local actions.

• If $\text{act} \in \text{upacts}(M_i)$, $s_i \xrightarrow{\text{act}_{i}} s_{0i}$, and $s_{R} \xrightarrow{\text{act}_{R}} s_{0R}$, then $(s_i, s_{R}) \xrightarrow{\text{act}_q} (s_{0i}, s_{0R})$, for all $j \in \{1, \ldots, n\}$. Intuitively, after synchronising with $R$, a component $M_i$ will reset and can release control to another module.

\[ L_{sq}(s_i, s_R) = L_R(s_i) \cup L_R(s_{R}), \text{ for each } (s_i, s_R) \in S_{sq}. \]

We call $S\mathcal{Q}^u(G)$ the Unreduced Sum-of-squares Product of $G$.

We say that a state $s$ of $G$ is locked iff there is no run $\rho \in \text{Runs}(G, s)$ s.t. $\rho = s^0 \text{act}^0 s^1 \text{act}^1 \ldots$ with $\text{act}^i \in \text{acts}(R)$, where $s^0 = s$, for some $i \in \mathbb{N}$. Observe that from a point of view of the root, a locked state is in a full deadlock. The set of locked states of an LTS can be computed in polynomial time using either a model checker or conventional graph algorithms.

**Definition 5 (Sum-of-squares Product).** We call the Sum-of-squares Product $S\mathcal{Q}(G)$ of $G$ the result of removing all the locked states from $S\mathcal{Q}^u(G)$ and restricting the relevant transition and labelling functions.

![Diagram](image-url)
Example 3. Fig. 2 presents the unreduced sum-of-squares product \( SQ^u(G_x) \) for the small tree \( ST \) from Example 1. The locked states are coloured red. Fig. 3 displays the sum-of-squares product \( SQ(G_x) \) of the topology. Note the similarity of the model to the root of Fig. 1 that reveals that the children do not restrict the root’s freedom.

As shown in the above example, it is possible that the size of the state space of an (unreduced) sum-of-squares product of a live tree \( ST \) is equal to or greater than the size of the state space of \( G \). On the other hand, the size of a representation of a state will be smaller in (unreduced) sum-of-squares product, as it records only local states of at most two components of the network. However, in less degenerate cases than our toy model we can expect significant reductions. In particular, if a two-level tree \( ST \) contains \( n \) components, where the state space of each is of size \( m \), then the size of its asynchronous product can reach \( m^n \). In contrast, the size of the (unreduced) sum-of-squares product of such topology is at most \((n-1) \cdot m^2\). The structure of the sum-of-squares is similar to the structure of the root. This construction preserves reachability, but not the \( EG \) modality of \( CTL \), as shown in Proposition 1.

**Theorem 1 (Sum-of-squares Preserves Reachability).** Let \( G \) be a live two-level tree \( ST \). For each \( p \in PV \) \( G \models EFp \) iff \( SQ(G) \models EFp \).

**Proof.** Recall that we assume \( Net = \{ R, M_1, \ldots, M_n \} \) with root \( R \) and children \( \{ M_i \}_{i=1}^n \). Let \( G \models EFp \) and \( \rho = s^0 \text{act}^0s^1 \text{act}^1 \ldots \) be a run of \( G \) s.t. \( p \in L(s_i) \) for some \( i \in N \). Now, \( \rho \) can be represented as \( \rho = \alpha_1 F_1 \alpha_2 F_2 \ldots \), where for each \( i \in N \) there exist \( j, k \in N \) such that \( \alpha_i = s^j \text{act}^j \ldots s^k \text{act}^k s^{k+1} \) and \( \text{act}^j, \ldots, \text{act}^k \in \text{locacts}(R) \cup \bigcup_{i=1}^n \text{locacts}(M_i) \) and \( F_i \in \text{downacts}(R) \). Actions are never synchronised between children, thus it can be proven by induction on the length of the run that the actions in \( \rho \) can be reordered to obtain a run \( \rho' \in \text{Runs}(G, s^0) \) that can be represented as \( \rho' = \alpha'_1 F_1 \alpha'_2 F_2 \ldots \), such that:

1. For any \( i \in N \) there exist \( j, k \in N \) such that \( \alpha'_i = s^{j'} \text{act}^{j'} \ldots s^{k'} \text{act}^{k'} s^{k'+1} \) and \( \text{act}^{j'}, \ldots, \text{act}^{k'} \in \text{locacts}(R) \cup \text{locacts}(\text{snd}(F_i)) \).
2. For each $i \in \mathbb{N}$ and $1 \leq j \leq n$ we have $\alpha_i \downarrow (R, \text{snd}(F_i)) = \alpha'_i \downarrow (R, \text{snd}(F_i))$.
3. For each $s'^j$ in $\alpha_i$, if 0 is the coordinate of root and $k$ is the coordinate of $\text{snd}(F_i)$, then $s'^j = (s_0, s_1^0, \ldots, s_{k-1}^0, s_k^0, s_{k+1}^0, \ldots)$ for some $s_0 \in \text{states}(R)$, $s_k \in \text{states}(\text{snd}(F_i))$.

Intuitively, $\rho'$ is built from $\rho$ in such a way that firstly only the root and the component that synchronises with the root over $F_1$ are allowed to execute their local actions while all the other components stay in their initial states; then $F_1$ is fired; and then this scheme is repeated for $F_2, F_3$, etc. We can now project $\rho'$ on spaces of squares of the root and components active in a given interval, to obtain $\rho'' = \alpha'_i \downarrow (R, \text{snd}(F_1))F_1 \alpha'_i \downarrow (R, \text{snd}(F_2))F_2 \ldots$ As $\rho'' \in \mathcal{SQ}(G)$ and it can be observed that $\rho''$ visits each local state that appears along $\rho$, this part of the proof is concluded.

Let $\mathcal{SQ}(G) \models EFp$ and $\rho \in \text{Runs}(\mathcal{SQ}(G))$ visit a state labelled with $p$. Now, it suffices to replace in $\rho$ each state $(s_k, s_0)$ that belongs to the square $M_k \times R$ with the global state $(s_0, s_1^0, \ldots, s_{k-1}^0, s_k, s_{k+1}^0, \ldots)$ of $G$. The result of this substitution is a run of $G$ that visits $p$.

**Proposition 1 (Sum-of-squares Does Not Preserve EG).** There exists a live two-level tree $\mathcal{ST} \ G$ s.t. for some $p \in \mathcal{PV} \ G \models EGp$ and $\mathcal{SQ}(G) \not\models EGp$.

**Proof.** Consider the tree $\mathcal{ST} \ G'_y$ in Fig. 4. Here, we have $G'_y \models EGp$, but each path $\rho$ along which $p$ holds globally, starts with $M_1^q$ executing $\tau$, followed by $M_2^q$ executing $\tau$ and, consecutively, $\text{chooseR}$. Thus, it is not possible to partition $\rho$ into intervals where one child executes local actions until synchronisation with the root and possible release of control to another child. Hence, $\mathcal{SQ}(G'_y) \not\models EGp$.

### 3.2 Adaptation for Any Tree Height

It is rather straightforward to adapt the sum-of-squares of a subtree to allow for synchronisation of the root with its parent. By $\text{cmpl}(\mathcal{SQ}(G))$ we denote the result of replacing in $\mathcal{SQ}(G)$ every transition $(s, act, s')$, where $act \in \text{upacts}(R)$ with $(s, act, s_{sq})$. Note that $\text{cmpl}(\mathcal{SQ}(G))$ is a live-reset tree $\mathcal{ST}$. 
Example 4. To obtain $\text{cmpl}(\mathcal{SQ}(\mathcal{G}))$ for the sum-of-squares product from Fig. 3 move the targets of the looped beep transitions to $s_{sq}'$.

We are now ready to provide the algorithm for reducing any live tree $\mathcal{ST}$ to a single component while preserving reachability.

**Alg. 1 reduceNet($\mathcal{G}$)**

**Input:** live-reset tree sync. topology $\mathcal{G}$

**Output:** LTS $\mathcal{M}$ s.t. $\mathcal{G} \models EFp$ iff $\mathcal{M} \models EFp$.

1: if $|\text{mods}(\mathcal{G})| = 1$ then
2: return $\mathcal{G}$ (* $\mathcal{G}$ is a leaf *)
3: end if
4: let redChldn := $\emptyset$
5: for $\text{child} \in \text{children(root(}\mathcal{G})\text{)}$ do
6: redChldn.append(reduceNet($\mathcal{G}_{\text{child}}$))
7: end for
8: let $\mathcal{G}' := \{\text{root(}\mathcal{G})\} \cup \text{redChldn}$
9: return $\text{cmpl}(\mathcal{SQ}(\mathcal{G}'))$

Algorithm 1 applies the two-level reduction $\text{cmpl}(\mathcal{SQ}(\cdot))$ to all the nodes of the $\mathcal{ST}$, in a bottom-up manner. Its soundness and correctness is expressed by the following theorem.

**Theorem 2 (reduceNet($\mathcal{G}$) Preserves Reachability).** Let $\mathcal{G}$ be a live tree $\mathcal{ST}$. For each $p \in PV \mathcal{G} \models EFp$ iff reduceNet($\mathcal{G}$) $\models EFp$.

**Proof.** (Sketch) The proof follows via induction on the height of the tree $\mathcal{G}$. As we have Theorem 1, it suffices to prove that $\text{cmpl}(\mathcal{SQ}(\mathcal{G}))$ preserves reachability for any two-level live $\mathcal{ST} \mathcal{G}$. This, however, can be done in a way very similar to the proof of Theorem 1 and is omitted.

4 Conclusion

In this paper we have outlined how to simplify large tree networks of automata that reset after synchronising with their parents. It is shown that the reduction preserves a certain form of reachability, but it does not preserve safety. While the procedure is quite fast and effective, it has several limitations. Firstly, it preserves reachability of labelings, but not their conjunctions; namely, it is not guaranteed that reduceNet($\mathcal{G}$) $\models EF(p \land q)$ iff $\mathcal{G} \models EF(p \land q)$. Secondly, we would like to relax the assumption that all the components are live-reset automata. It is not difficult to see how to adapt the original construction to the general case. To this end it suffices to extend the sum-of-squares product with an explicit model of the memory of last synchronisations. Interacting modules can then use this memory to register the return states, i.e. the locations entered after synchronising action.
Thus, a synchronising step between a root and one of its children would become a process consisting of the following four steps: (1) perform joint synchronising transition; (2) record the target locations; (3) make a non-deterministic selection of a child and read from the memory its return state; (4) continue the execution of the pair of the root and the new child. This construction, however, can hinder the expected reduction due to the size of the memory component. Finally, it is possible that the assumption of tree-like communication between the components is too strong for any real-life applications. Thus it should be investigated if the proposed procedures can be easily extended to other topologies.

We plan to address these limitations in future work.

References

1. “LTR,” https://github.com/MichalKnapik/automata-net-reduction-tool, 2020.
2. G. J. Holzmann, The SPIN Model Checker - primer and reference manual. Addison-Wesley, 2004.
3. G. Behrmann, A. David, K. G. Larsen, J. Håkansson, P. Pettersson, W. Yi, and M. Hendriks. “UPPAAL 4.0,” in Third International Conference on the Quantitative Evaluation of Systems (QEST 2006), 11-14 September 2006, Riverside, California, USA. IEEE Computer Society, 2006, pp. 125–126.
4. É. André, L. Fribourg, U. Kühne, and R. Soulat, “IMITATOR 2.5: A tool for analyzing robustness in scheduling problems,” in FM 2012, ser. LNCS, vol. 7436. Springer, 2012, pp. 33–36.
5. C. Baier and J. Katoen, Principles of model checking. MIT Press, 2008.