

Control of safe ordinary Petri nets using unfolding*

Alessandro Giua (*), Xiaolan Xie (**)

(*) Dip. di Ingegneria Elettrica ed Elettronica, Università di Cagliari,

Piazza d'Armi, 09123 Cagliari, Italy.

Email: giua@diee.unica.it

(**) Ecole Nationale Supérieure des Mines de Saint-Etienne

158 cours Fauriel, 42023 Saint-Etienne cedex 2, France.

Email: xie@emse.fr

Abstract

In this paper we deal with the problem of controlling a safe place/transition net so as to avoid a set of forbidden markings \mathcal{F} . We say that a given set of markings has property REACH if it is closed under the reachability operator. We assume that all transitions of the net are controllable and that the set of forbidden markings \mathcal{F} has the property REACH.

The technique of unfolding is used to design a maximally permissive supervisor to solve this control problem. The supervisor takes the form of a set of control places to be added to the unfolding of the original net.

The approach is also extended to the problem of preventing a larger set \mathcal{F}_I of impending forbidden marking. This is a superset of the forbidden markings that also includes all those markings from which - unless the supervisor blocks the plant - a marking in \mathcal{F} is inevitably reached in a finite number of steps.

Finally, we consider the particular case in which the control objective is that of designing a maximally permissive supervisor for deadlock avoidance and we show that in this particular case our procedure can be efficiently implemented by means of linear algebraic techniques.

1 Introduction

In this paper we show how Petri net *unfolding* can be used to design a maximally permissive supervisor to avoid a set of forbidden markings \mathcal{F} . We assume that all transitions are controllable and that the set \mathcal{F} has property REACH, i.e., any marking reached from a marking in \mathcal{F} is also forbidden.

*Published as: A. Giua, X. Xie, Control of safe ordinary Petri nets using unfolding, *Discrete Event Dynamic Systems*, Vol. 15, No. 4, pp. 349–373, Dec 2005. The original publication is available at www.springerlink.com.

1.1 Relevant literature

The use of *partial order methods* for the efficient verification of *concurrent systems* is a technique that has been used by several authors in the last 10-15 years. In particular, a Petri net is a natural model for this approach because it has primitives to explicitly capture the notion of precedence and independence between events.

The interleaving of concurrent sequences often leads to the well-known problem of *state space explosion* that hinders the applicability of all those Petri net analysis techniques, e.g., the reachability graph, that are based on the exhaustive search over the set of reachable markings. However, the sets of states introduced by concurrency are for the most part intermediate markings that are irrelevant to determine the properties of the system: what matters is the unique marking reached by the firing of all these concurrent sequences.

As Valmari (1994) has lucidly explained, this fact has motivated research along at least two different lines.

- A first approach is to let one (or at least as few as possible) interleaving represent all its equivalent interleavings: the notion of *stubborn set* (Valmari, 1991) and *persistent set* (Godefroid, 1996) is inspired by this idea.
- A second approach consists in replacing the reachability graph by a net structure which captures the concurrent executions, and does not explicitly show individual interleavings. This technique is based on the *unfolding* of a (bounded) Petri net into an *occurrence net*. A finite prefix of the unfolding can be used to characterize the set of all reachable markings without having to enumerate them (McMillan, 1995; Esparza *et al.*, 2002). Recently, this approach has also been extended to unbounded nets (Neumair, 2002). Note that the occurrence net is much simpler than the original Petri net and can usually be validated using structural analysis.

Although these two types of techniques have proved to be a powerful instrument in the *verification* of concurrent systems, the application of these techniques to the *control* of discrete event systems has not received a lot of attention. We recall here some contributions in this area.

Hellgren *et al.* (1999) have used persistent sets to design supervisors for deadlock avoidance.

Observability and diagnosis are closely related to control: Aghasaryan *et al.*, (1998) were the first to use unfolding for fault detection and diagnosis in distributed systems. This approach has also been extended in two subsequent papers by Benveniste *et. al* (2003a; 2003b).

Recently, in a series of papers He and Lemmon (2000; 2002) have presented an original approach based on unfolding for liveness verification and enforcing. However we have shown (Xie and Giua, 2004) that some key results of these papers are incorrect. As a result, although we still strongly believe that unfolding is an interesting and potentially fruitful technique for Petri net control, the applicability of unfolding for Petri net supervision is still an open issue.

Part of the material presented in this paper has also been presented in (Giua and Xie, 2004; Giua and Xie, 2005).

1.2 Contribution

In the paper we consider discrete event systems modelled by *safe place/transition nets*. The control problem we consider can be framed within the theory of *supervisory control* (Ramadge and Wonham, 1989). In particular, we consider a control specification that requires avoiding a set of *forbidden markings* \mathcal{F} . In the current state of investigation, we assume that all transitions are controllable, i.e., they can be disabled by a controlling agent called supervisor that must enforce the specification.

We use a set of finite prefixes of the unfolding, that we call *order 1 unfolding* and *order 2 unfolding*, to characterize the reachability set of the original net.

We restrict our attention to a special class of forbidden markings specifications that satisfy what we call *property REACH*: once a forbidden marking is reached, all markings reachable from it will also be forbidden. This has a nice advantage over the unfolding structure: if a configuration (i.e., a set of transition firings) is forbidden, any larger configuration should also be forbidden. We show that in this case a simple control structure - that consists of a set of places to be added to the order 1 unfolding - can be used to implement a maximally permissive control policy that enforces the specification.

In many control problems it is necessary not only to ensure a safe behavior of the controlled system — e.g., avoiding a set of forbidden markings — but to guarantee some liveness properties as well. A usual liveness requirement is that the supervisor should not deadlock the system. To address this problem, we consider the problem of preventing the larger set \mathcal{F}_I of *impending forbidden markings*. This is a superset of the forbidden markings that also includes all those markings from which - unless the supervisor blocks the plant - a marking in \mathcal{F} is inevitably reached in a finite number of steps. In this case, we use the larger order 2 unfolding to compute a set of control places that, added to the order 1 unfolding, can be used to implement a maximally permissive control policy for this problem.

The approach we present in the paper requires in general an exhaustive enumeration of the set of forbidden markings. It has however the advantage of allowing one to construct a maximally permissive supervisor in the form of a *controlled occurrence net* (i.e., an occurrence net with the addition of control places) using a procedure which does not require the exhaustive enumeration of the set of markings of the plant. The closed loop system in this approach can also be represented by this controlled occurrence net.

In many cases it is also possible to avoid the exhaustive enumeration of the set of forbidden markings and solve the control problem by means of structural analysis. Two examples of this type are discussed in the paper.

A first example is the case in which the set of forbidden markings satisfies a set of linear inequalities. In this case it is possible to check if the set has property REACH and to design the maximally permissive controller by integer programming analysis.

A second, and more meaningful, example is the case in which the control objective is that of designing a *deadlock prevention control policy*. This is a problem that has received a lot of attention in the literature (Ezpeleta *et al.*, 1995; Chu and Xie, 1997; Park and Reveliotis, 2001) and the last section of the paper is devoted to this problem. If the set of forbidden markings \mathcal{F} consists of the set of deadlock markings of the original net, the maximally permissive nonblocking control policy consists exactly in the prevention of the set \mathcal{F}_I and we show in the paper that the computation of the supervisor can be carried out by

structural analysis.

The following two remarks should clarify the contribution of this paper.

- Our approach deals with a particular class of *marking specifications* and not with more general *language specifications* usually considered in supervisory control theory (Ramadge and Wonham, 1989). We believe that it may be possible to extend this approach to the latter type of specifications, but this issue will be the object of future research.
- The computational advantage of our approach lies in the fact that it does not require to exhaustively generate the whole state space of the plant (i.e., its reachability set), because it works on the structure of the unfolding net. This advantage can only be gauged in qualitative terms. In fact, if the plant is composed by several subsystems with a high degree of concurrency, then the size of the unfolding is much smaller than the size of the reachability set. If, on the contrary, the behavior of the plant does not contain many interleavings of concurrent sequences, then the advantages of working on the unfolding is less significant.

This paper is structured as follows. In Section 2 we recall the standard notation on Petri nets. Section 3 contains an informal presentation of unfolding and we also define the finite prefixes of the unfolding that will be used in this paper. In Section 4 we discuss a special class of forbidden markings specification having property REACH that will be considered in the following. In Section 5 we show how the order 1 unfolding can be used to design a maximally permissive supervisor for \mathcal{F} . In Section 6 we discuss the problem of preventing the set \mathcal{F}_I . Finally, in Section 7 we show how a maximally permissive supervisor for deadlock avoidance can be efficiently designed using linear algebraic techniques.

2 Background on Petri nets

In this section we recall the formalism used in the paper. A more detailed introduction to Petri nets can be found in (Murata, 1989).

The Petri net model considered in this paper is an *ordinary Place/Transition net* (P/T net) denoted $N = (P, T, F)$, where P is a set of m places; T is a set of n transitions; $F \subseteq (P \times T) \cup (T \times P)$ is the flow function that specifies the arcs from places to transitions and from transitions to places.

The *incidence matrix* C of a net is an $m \times n$ matrix such that $C(p, t) = 1$ if $(t, p) \in F$ and $(p, t) \notin F$, $C(p, t) = -1$ if $(p, t) \in F$ and $(t, p) \notin F$, else $C(p, t) = 0$.

The *preset* and *postset* of a node $x \in P \cup T$ are denoted $\bullet x \triangleq \{x' \mid (x', x) \in F\}$ and $x^\bullet \triangleq \{x' \mid (x, x') \in F\}$ while $\bullet x^\bullet = \bullet x \cup x^\bullet$. Node x is a *source* (resp., *sink*) if $\bullet x = \emptyset$ (resp.; $x^\bullet = \emptyset$).

Given two nodes $x, x' \in P \cup T$ we define the following relations.

- Node x *precedes* x' (denoted $x \preceq x'$) if there exists a directed path from x to x' . If we require that the path has length greater than zero (i.e., $x \neq x'$) we write $x \prec x'$.
- Nodes x and x' are in *conflict* (denoted $x \# x'$) if there exist two different transitions $t, t' \in T$ such that: $t \preceq x, t' \preceq x', \bullet t \cap \bullet t' \neq \emptyset$. In this case, in fact transitions t and t' are in conflict because they

have a common input place, and the conflict propagates to all nodes following them. A node x is in *self-conflict* if $x\#x$ holds.

- Nodes x and x' are *concurrent* (denoted $x \approx x'$) if neither $x \preceq x'$, nor $x' \preceq x$, nor $x\#x'$ hold.

Note that given two nodes x and x' it may hold that: ($x \prec x'$ and $x' \prec x$ and $x\#x'$).

A *marking* is a vector $M : P \rightarrow \mathbb{N}$ that assigns to each place of a P/T net a non-negative integer number of tokens, represented by black dots. We denote $M(p)$ the marking of place p . A P/T *system* or *net system* $\langle N, M_0 \rangle$ is a net N with an initial marking M_0 .

A transition t is *enabled* at M iff $M(p) > 0$ for all $p \in \bullet t$. If t is enabled, it may *fire* yielding the marking $M' = M + C(\cdot, t)$. We write $M[\sigma]$ to denote that the sequence of transitions $\sigma = t_{j_1} \cdots t_{j_k}$ is enabled at M , and we write $M[\sigma]M'$ to denote that the firing of σ yields M' . We associate to a sequence σ a *firing vector* $X : T \rightarrow \mathbb{N}$ such that $X(t) = k$ if transition t appears k times in σ .

A marking M said to be *dead* if no transition is enabled at M .

A marking M is *reachable* in $\langle N, M_0 \rangle$ iff there exists a firing sequence σ such that $M_0[\sigma]M$. The set of all markings reachable from M_0 defines the *reachability set* of $\langle N, M_0 \rangle$ and is denoted $R(N, M_0)$.

We denote $PR(N, M_0)$ the *potentially reachable set*, i.e., the set of all markings $M \in \mathbb{N}^m$ for which there exists a vector $X \in \mathbb{N}^n$ that satisfies the *state equation* $M = M_0 + C \cdot X$, i.e., $PR(N, M_0) \triangleq \{M \in \mathbb{N}^m \mid \exists X \in \mathbb{N}^n : M = M_0 + C \cdot X\}$. It holds that $R(N, M_0) \subseteq PR(N, M_0)$.

A place p is *k-bounded* if for all $M \in R(N, M_0)$ it holds $M(p) \leq k$. A *1-bounded* place is called *safe*. A net system $\langle N, M_0 \rangle$ is said *k-bounded* (resp., *safe*) if all its places are *k-bounded* (resp., *safe*). A marking M of a safe net system is a binary vector and can also be seen as a set of places $M = \{p \in P \mid M(p) = 1\}$.

In the rest of the paper for sake of simplicity we will consider only safe net systems. However, the results presented in this paper can also be extended to arbitrary bounded nets in a straightforward manner.

3 Unfolding

In this section we informally recall how it is possible, given a safe net system $\langle N, M_0 \rangle$, to *unfold* it by constructing a *labelled occurrence net* $\tilde{N}(M_0)$. This occurrence net, that is also commonly called the *unfolding of* $\langle N, M_0 \rangle$, has a structure that depends both on N and on M_0 . A formal description of the unfolding procedure requires a long and tedious series of definitions: we prefer to present the key concepts here. Any of the references (McMillan, 1995; Esparza *et al.*, 2002; He and Lemmon, 2002; Benveniste *et al.*, 2003b) contains a more comprehensive and accurate discussion.

An *occurrence net* is an ordinary P/T net with a special structure:

- starting from any node, all backward paths are finite, i.e., eventually they reach a source node;
- each place has at most one input arc;
- no node is in self-conflict.

It is easy to show that in an occurrence net the two relations of precedence and conflict are mutually exclusive and that the precedence relation is a partial order. Thus, for any two distinct nodes x and x' one and only one of the following conditions holds: $x \prec x'$, or $x' \prec x$, or $x \# x'$, or $x \approx x'$. The net shown in Figure 2 is an example of a finite occurrence net.

To the unfolding $\tilde{N}(M_0) = (\tilde{P}, \tilde{T}, \tilde{F})$ is also associated a *labelling function* $\ell : (\tilde{P} \rightarrow P) \cup (\tilde{T} \rightarrow T)$ that maps each node of the unfolding into a node of the original net N . Note that usually a node p or t of N may correspond to more than one node of the unfolding, i.e., $\ell^{-1}(p) \subset \tilde{P}$ and $\ell^{-1}(t) \subset \tilde{T}$.

The labelling function can also map set of nodes into set of nodes. In particular, in the following procedure given a set of places $P' \subseteq P$ of the original net, we write $P' = \hat{\ell}(\tilde{P}')$ to denote that the set of places \tilde{P}' of the unfolding has the same cardinality of P' and $P' = \left\{ p \in P \mid \tilde{p} \in \tilde{P}', p = \ell(\tilde{p}) \right\}$, hence each place of \tilde{P}' maps into a place of P' but no two places in \tilde{P}' map into the same place of P' .

Procedure 1. (Unfolding of a safe net system $\langle N, M_0 \rangle$ into an occurrence net $\tilde{N}(M_0)$)

1. Add to the unfolding a set of source places \tilde{P}_0 with $\hat{\ell}(\tilde{P}_0) = \{p \in P \mid M_0(p) = 1\}$.
2. Let $i := 0$.
3. Let $\tilde{P}_{\text{exp}} := \tilde{P}_i$
4. If $\tilde{P}_i = \emptyset$ then STOP.
5. Let $i := i + 1$.
6. Let $\tilde{P}_i := \emptyset$.
7. For all transitions $t \in T$

For all sets of places $\tilde{P}' \subseteq (\tilde{P}_{\text{exp}} \setminus \tilde{P}_i)$ such that the following three conditions are all verified:

- $\hat{\ell}(\tilde{P}') = \bullet t$,
- all places in \tilde{P}' are concurrent,
- $\tilde{P}' \cap \tilde{P}_{i-1} \neq \emptyset$,

- (a) Add to the unfolding a new transition \tilde{t} with $\hat{\ell}(\tilde{t}) = t$.
- (b) Add to the unfolding a set of new places \tilde{P}'' with $\hat{\ell}(\tilde{P}'') = t \bullet$.
- (c) Add an arc from each place in \tilde{P}' to \tilde{t} .
- (d) Add an arc from \tilde{t} to each place in \tilde{P}'' .
- (e) Let $\tilde{P}_i := \tilde{P}_i \cup \tilde{P}''$.
- (f) Let $\tilde{P}_{\text{exp}} := \tilde{P}_{\text{exp}} \cup \tilde{P}''$.

8. Goto 4. ■

In the procedure at step 1 we add to the unfolding a copy of each place of the original net marked by the initial marking: all places in this set \tilde{P}_0 are ranged on the tier 0 and represent the source nodes of the occurrence net. The index i initialized at step 2 denotes the tier on which the places of each set \tilde{P}_i are ranged.

The set \tilde{P}_{exp} initialized at step 3 keeps track of the places that can be used to expand the unfolding. Strictly speaking, in this version of the procedure this set needs not be defined because it always coincides with the set of places in the unfolding. However we will need it when the procedure is modified to construct a finite prefix (that we call order 1 unfolding) as explained in the following.

Each time a new tier is added we check for all transitions t of the original net if there exists in the unfolding a set \tilde{P}' with the following properties:

- it is a copy of the set of input places of t , hence a marking that marks in the unfolding all places in \tilde{P}' corresponds to a marking of the original net that enables t ;
- all places in \tilde{P}' are concurrent, hence they can simultaneously be marked (note that the safeness of the original net ensures that two places on the unfolding with the same label cannot be concurrent);
- at least one place in \tilde{P}' belongs to the lastly added tier, so that the marking of the unfolding that marks all these places has not been considered in the previous steps.

For all such sets \tilde{P}' we add to the net a new copy of transition t , a new copy of all its output places \tilde{P}'' and the relative arcs.

The procedure given above is not an algorithm because it is not guaranteed to halt in a finite number the steps. In fact the unfolding of a net that admits repetitive sequences is infinite.

Example 2. *The net in Figure 2 is a finite prefix of the unfolding of the safe net¹ in Figure 1, constructed using the previous procedure. A place \tilde{p} of the unfolding such that $\ell(\tilde{p}) = p_k$ is labelled k, k', \dots . A transition \tilde{t} of the unfolding such that $\ell(\tilde{t}) = t_k$ is labelled k, k', \dots ■*

Note that we can consider an unfolding both as a net and as a marked net where the initial marking assigns to each source place in \tilde{P}_0 a token, so we need not specify its initial marking and simply write $R(\tilde{N}(M_0))$ to denote its reachability set.

The unfolding is a safe net so we can represent a marking with the set of non-empty place: we write $\tilde{M}_0 = \tilde{P}_0$ and in general $\tilde{M} = \{\tilde{p} \in \tilde{P} \mid \tilde{M}(\tilde{p}) = 1\}$. It is also possible to apply the mapping $\hat{\ell}$ to markings.

Definition 3. *To each marking \tilde{M} of the unfolding corresponds a marking of the original net $M = \hat{\ell}(\tilde{M}) \triangleq \{p \in P \mid p = \ell(\tilde{p}), \tilde{p} \in \tilde{M}\}$. This leads to an equivalence relation among markings in $R(\tilde{N}(M_0))$ and if $\hat{\ell}(\tilde{M}) = \hat{\ell}(\tilde{M}')$ we write $\tilde{M} =_P \tilde{M}'$.*

A firing vector \tilde{X} of the unfolding is a binary vector that can also be seen as a set of transitions $\tilde{X} = \{\tilde{t} \in \tilde{T} \mid \tilde{X}(\tilde{t}) = 1\}$.

Definition 4. *Given a transition $\tilde{t} \in \tilde{T}$, the minimal firing vector of the unfolding that contains it is called a local configuration; it can be show that this vector is unique and we denote it $[\tilde{t}]$. The marking reached firing configuration \tilde{X} (resp., $[\tilde{t}]$) will be denoted $\tilde{M}(\tilde{X})$ (resp., $\tilde{M}([\tilde{t}])$).*

It is also clear that each marking \tilde{M} reachable in $\tilde{N}(M_0)$ corresponds to a unique configuration in $\tilde{N}(M_0)$ (the unfolding net is acyclic) that we sometimes denote $\text{conf}(\tilde{M})$.

¹In the net in Figure 1 an arc with a double arrow between a place p and a transition t denotes a self-loop, i.e., $(p, t) \in F$ and $(t, p) \in F$.

Given a net system $\langle N, M_0 \rangle$, McMillan (1995) presented a technique to construct a finite prefix of its unfolding. Following Lemmon and He (2002), we consider a slightly different construction of the finite prefix.

Definition 5 (Order 1 unfolding). *The order 1 unfolding, denoted $\tilde{N}_1(M_0)$, is a finite prefix of the unfolding obtained by Procedure 1 stopping the construction of the unfolding when we reach a cut-off transition \tilde{t} , i.e., a transition such that:*

- *EITHER the firing of the local configuration of \tilde{t} brings back to the initial marking, i.e., $\tilde{M}([\tilde{t}]) =_P \tilde{M}_0$;*
- *OR there exists another transition \tilde{t}' with the following properties:*
 - (a) *\tilde{t}' has a smaller configuration than \tilde{t} : $[\tilde{t}'] \subset [\tilde{t}]$;*
 - (b) *the markings reached firing the two configurations are equivalent, i.e., $\tilde{M}([\tilde{t}']) =_P \tilde{M}([\tilde{t}])$.*

In the following we call \tilde{t}' the mirror transition of \tilde{t} in $\tilde{N}_1(M_0)$.

It should be noted that what we call order 1 unfolding is a net slightly larger than McMillan finite prefix, because our condition (a) is stronger: McMillan requires only that $\text{card}([\tilde{t}']) < \text{card}([\tilde{t}])$.

Algorithm 6. *The order 1 unfolding can be constructed using a modified version of Procedure 1 where the instruction 7.(f) is changed to*

7.(f') If t is not a cut-off transition, then let $\tilde{P}_{\text{exp}} := \tilde{P}_{\text{exp}} \cup \tilde{P}''$.

In this case when a cut-off transitions is added to the unfolding, a copy of its output places is also added but they will not be used to expand the unfolding any further. With this small change, Procedure 1 always stops in a finite number of steps if the original system is safe (and can now be called an algorithm).

Example 7. Consider the net shown in Figure 1. Its order 1 unfolding is shown in Figure 2 (the subnet in darker color). Note that we have also added to the unfolding the cut-off transitions and their output places: the cut-off transitions of the order 1 unfolding are the thick transitions \tilde{t}_2 (on tier 3) and \tilde{t}_6 (on tier 4). Transition \tilde{t}_5 (on tier 2) is not a cut-off transition: after its firing the unfolding cannot proceed because a deadlock is reached. ■

The following result follows from an original result presented by McMillan (1995).

Lemma 8. *The image through the labelling function of the reachability set of the order 1 unfolding $\tilde{N}_1(M_0)$ is the reachability set of the original system, i.e.,*

$$R(N, M_0) = \ell(R(\tilde{N}_1(M_0))) \triangleq \left\{ M \in \mathbb{N}^m \mid M = \ell(\tilde{M}), \tilde{M} \in R(\tilde{N}_1(M_0)) \right\}.$$

Proof. McMillan showed this result holds for the total unfolding and also for the finite prefix. Since $\tilde{N}_1(M_0)$ is larger than McMillan finite prefix, the result follows immediately. □

We can also define a larger finite prefix of the unfolding.

Definition 9 (Order 2 unfolding). *Once constructed $\tilde{N}_1(M_0)$, assume we continue the unfolding until we reach a transition \tilde{t} such that there exist a transition \tilde{t}' with the following properties:*

- (a) *either \tilde{t}' does not belong to $\tilde{N}_1(M_0)$ or it is a cut-off transition of $\tilde{N}_1(M_0)$;*

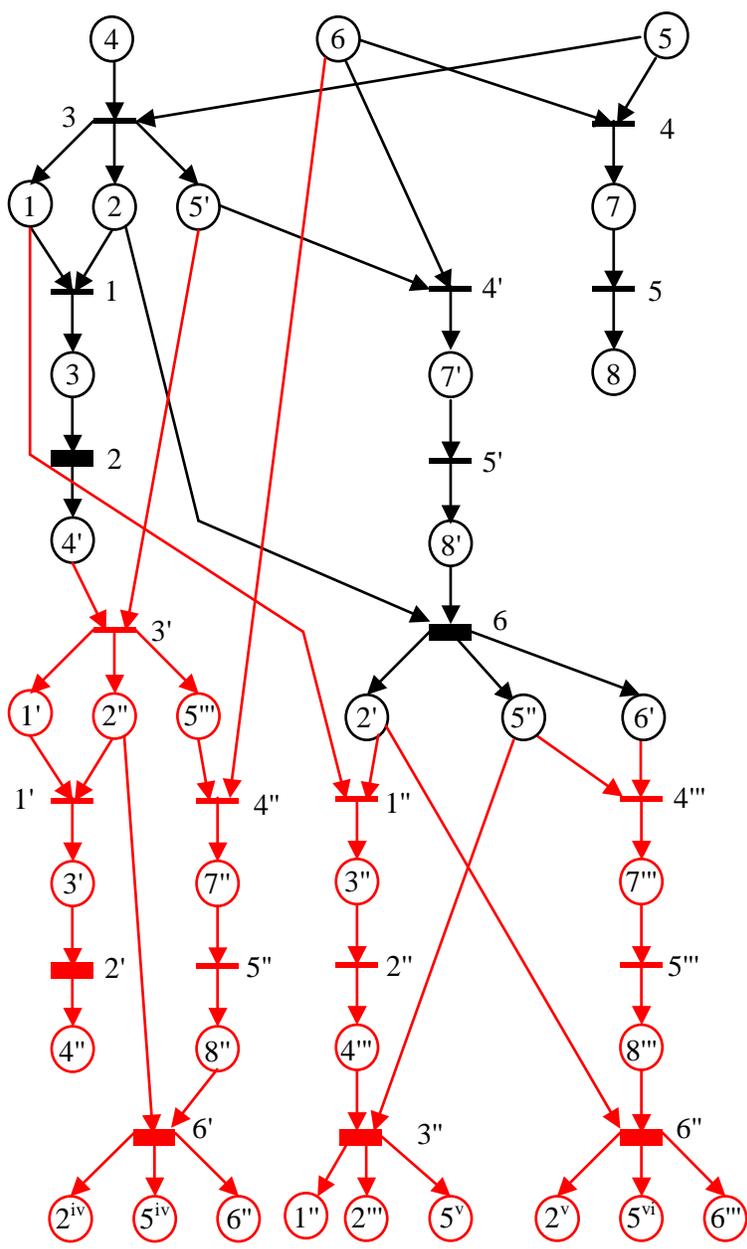


Figure 2: The order 1 unfolding (subnet in darker color) and the order 2 unfolding (the complete net) of the net in Figure 1.

4 A special class of forbidden markings problem

We consider a control problem where the set of forbidden markings \mathcal{F} has a special structure.

Definition 13. A set $\mathcal{F} \subseteq R(N, M_0)$ has property REACH wrt a net system (N, M_0) if

$$M \in \mathcal{F} \text{ and } M' \in R(N, M) \Rightarrow M' \in \mathcal{F}.$$

Thus property REACH implies that the set is closed under the reachability operator.

Meaningful examples of sets that have property REACH are the following:

- the set of deadlock markings;
- the set of markings from which there exists no firing sequence containing a given transition;
- the set of markings that are not co-reachable, i.e., from which it is not possible to reach a given set of final markings;
- the set of markings from which the initial marking is not reachable, i.e., from which no control law can ensure reversibility;
- the set of markings from which there exists no firing sequence containing all transitions, i.e., from which no control law can ensure liveness.

An example of a class of sets that do not necessarily have property REACH is the set of forbidden markings corresponding to a *generalized mutual exclusion constraint*² (w, k) . The set of forbidden markings for such a constraint, where $w \in \mathbb{Z}^m$ and $k \in \mathbb{Z}$, is $\mathcal{F}(w, k) = \{M \in R(N, M_0) \mid w^T M > k\}$, and it is easy to find examples where the property REACH does not hold. As an example, in the net in Figure 1 consider the GMEC (w, k) with

$$w = [0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 1 \ 0]^T \quad \text{and} \quad k = 1.$$

This constraint forbids the set

$$\mathcal{F}(w, k) = \{M \in R(N, M_0) \mid M(p_3) + M(p_7) > 1\} = \{[0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 1 \ 0]^T\}$$

that does not have property REACH because from the unique marking in this set the firing of, say, t_2 leads to $[0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 1 \ 0]^T \notin \mathcal{F}$.

In the following we focus on the optimal control of \mathcal{F} . Property REACH will allow us to use unfolding to design optimal controllers, as we show in the following section. However, for some control problems it is not sufficient to prevent a net from reaching markings in a set \mathcal{F} but it is also necessary to prevent the set \mathcal{F}_I of markings that will inevitably lead to a marking in \mathcal{F} .

Definition 14. Given a set $\mathcal{F} \subseteq R(N, M_0)$ we define its impending set as

$$\mathcal{F}_I = \{M \in R(N, M_0) \mid (\exists k \in \mathbb{N}) : R_{\geq k}(N, M) \subseteq \mathcal{F}, R_{< k}(N, M) \cap D(N, M_0) \subseteq \mathcal{F}\},$$

where $R_{\geq k}(N, M)$ (resp., $R_{< k}(N, M)$) denotes the set of markings reachable from M with a firing sequence containing at least (resp., less than) k transitions, and $D(N, M_0)$ is the set of dead markings of the net system $\langle N, M_0 \rangle$. ■

² A generalized mutual exclusion constraint (or GMEC) is a linear constraint on the set of reachable markings as defined in (Giua *et al.*, 1992).

Thus, starting from a marking in \mathcal{F}_I any evolution of length k or more and any evolution of length less than k that cannot be continued leads to \mathcal{F} . Clearly if a marking in $\mathcal{F}_I \setminus \mathcal{F}$ is reached, the only means the supervisor has to prevent the plant from reaching a marking in \mathcal{F} is that of blocking it. Hence avoiding \mathcal{F}_I allows the supervisor to prevent \mathcal{F} without having to block the plant. Note that by definition $\mathcal{F} \subseteq \mathcal{F}_I$.

If the larger set \mathcal{F}_I must be avoided, property REACH is still preserved as shown in the following result.

Theorem 15. *If a set \mathcal{F} has property REACH, then the set \mathcal{F}_I has property REACH.*

Proof. Consider any marking $M \in \mathcal{F}_I$. From the definition, there exists an integer k such that $R_{\geq k}(N, M) \subseteq \mathcal{F}$, $R_{< k}(N, M) \cap D(N, M_0) \subseteq \mathcal{F}$. Consider any marking M' reachable from M such that $M' \in R_i(N, M)$. If $i \geq k$, then $M' \in \mathcal{F}$ and hence $M' \in \mathcal{F}_I$. If $i < k$, then $M' \in \mathcal{F}_I$ since $R_{\geq k-i}(N, M') \subseteq R_{\geq k}(N, M) \subseteq \mathcal{F}$, and $R_{< k-i}(N, M') \cap D(N, M_0) \subseteq R_{< k}(N, M) \cap D(N, M_0) \subseteq \mathcal{F}$. \square

The following result shows an important consequence of the property REACH on the structure of an unfolding net.

Theorem 16. *Given a set \mathcal{F} with property REACH and a marking \tilde{M} such that $\hat{\ell}(\tilde{M}) \in \mathcal{F}$, if \tilde{M} is reachable with configuration \tilde{X} , then any larger configuration $\tilde{X}' \geq \tilde{X}$ leads to a marking \tilde{M}' such that $\hat{\ell}(\tilde{M}') \in \mathcal{F}$.*

Proof. If \tilde{M} is reachable with configuration \tilde{X} , and \tilde{M}' is reachable with configuration \tilde{X}' then:

$$\tilde{M} = \tilde{M}_0 + \tilde{C}\tilde{X}, \quad \text{and} \quad \tilde{M}' = \tilde{M}_0 + \tilde{C}\tilde{X}'.$$

This implies

$$\tilde{M}' = \tilde{M} + \tilde{C}(\tilde{X}' - \tilde{X})$$

with $\tilde{X}' - \tilde{X} \in \mathbb{N}^{\tilde{n}}$ hence \tilde{M}' is reachable by \tilde{M} (by Corollary 12). Thus $\hat{\ell}(\tilde{M}')$ must be reachable from $\hat{\ell}(\tilde{M})$ and (by property REACH) it belongs to \mathcal{F} . \square

To conclude this section, let us comment how it may be possible to check if a given set \mathcal{F} has property REACH. In the general case this may be done with an exhaustive reachability analysis: this is always possible because the reachability set of a safe net is finite. However, there are some special cases in which more efficient techniques can be used.

As an example, assume that the set of forbidden markings is given as a linear set:

$$\mathcal{F} = \{M \in \mathbb{N}^m \mid AM \leq b\}, \tag{1}$$

where A is a matrix $r \times m$ and b is a vector $r \times 1$. In this case is also immediate to observe that the set of markings of the unfolding corresponding to forbidden markings are

$$\ell^{-1}(\mathcal{F}) = \{\tilde{M} \in \mathbb{N}^{\tilde{m}} \mid \tilde{A}\tilde{M} \leq b\},$$

where \tilde{A} is a suitable matrix $r \times \tilde{m}$ that can be constructed from A . In particular, assume that the matrix A has elements $A(i, p)$ where $i = 1, \dots, r$, denotes the generic constraint, and $p \in P$ is a place of the net. Then for all $i = 1, \dots, r$ and for all $\tilde{p} \in \ell^{-1}(p)$ it holds $\tilde{A}(i, \tilde{p}) = A(i, p)$.

Under this assumption the following theorem holds.

Theorem 17. *A forbidden set \mathcal{F} of the form given in eq. (1) has property REACH if and only if the following IPP*

$$\left\{ \begin{array}{ll} \min J = \sum_{i=1}^r z_i & \\ \text{s.t.} & \begin{array}{ll} \tilde{M} = \tilde{M}_0 + \tilde{C}_2 \tilde{X} & (a) \\ \tilde{A} \tilde{M} \leq b & (b) \\ \tilde{M}' = \tilde{M} + \tilde{C}_2 \tilde{X}' & (c) \\ zK + \tilde{A} \tilde{M}' - (b + 1_r) \geq 0 & (d) \end{array} \\ \tilde{M}, \tilde{M}' \in \mathbb{N}^{\tilde{m}}; \tilde{X}, \tilde{X}' \in \mathbb{N}^{\tilde{n}}; z \in \{0, 1\}^r. & \end{array} \right.$$

has optimal performance index $J^* < r$, where K is a sufficiently large number, 1_r is a $r \times 1$ vector of 1's, and \tilde{C}_2 is the incidence matrix of the order 2 unfolding.

Proof. The proof can then be established by the following facts: (1) each forbidden marking M has an equivalent marking \tilde{M} in the order 1 unfolding and (2) the set of markings reachable from \tilde{M} in the order 2 unfolding contains the set of markings reachable from M in the original net N (this result is formally proved later in Theorem 29).

Obviously $J^* < r$ if and only if there exist at least one i such that $z_i = 0$. This means that there exists a marking \tilde{M} that is reachable (a) and forbidden (b) from which a marking \tilde{M}' is reachable (c) and \tilde{M}' is not forbidden (d) because it does not satisfy the i -th equation that defines $\ell^{-1}(\mathcal{F})$. Note that the state equation can be used to characterize reachability in (a) and (c) thanks to Corollary 12. \square

5 Control policy for \mathcal{F}

Given a forbidden markings set \mathcal{F} with property REACH we now present a maximally permissive control policy ensuring that no marking in \mathcal{F} is reached. This control policy will be “implemented” in the unfolding net by places with output arcs and no input arcs.

For marking $\tilde{M} \in R(\tilde{N}(M_0))$ such that $\hat{\ell}(\tilde{M}) \in \mathcal{F}$ let \tilde{X} be the unique configuration that yields it.

Definition 18. *The set of control transitions of \tilde{M} is $\tilde{X}_c = \left\{ \tilde{t} \in \tilde{X} \mid \tilde{A} \tilde{t}' \in \tilde{X}, \tilde{t} \in [\tilde{t}'] \right\}$.*

In plain words, these are all transitions inputting into the places that belong to \tilde{M} and that do not precede any other such transition.

It is easy to prove that all these transitions are concurrent. In fact, since $\tilde{X}_c \subseteq \tilde{X}$ and \tilde{X} is a fireable sequence, no two transitions can be in conflict. Furthermore, all transitions preceding another one in the set \tilde{X} are removed by construction, thus we are left with only concurrent transitions in \tilde{X}_c .

We will use the following control structure to prevent reaching \tilde{M} .

Definition 19. *Given a marking \tilde{M} with set of control transitions \tilde{X}_c , the control place \tilde{p}_c for \tilde{M} is a new place initially marked with $|\tilde{X}_c| - 1$ tokens and with an arc going to each transition in \tilde{X}_c . The incidence matrix of the control place is $\tilde{C}(\tilde{p}_c, \tilde{t}) = -1$ if $\tilde{t} \in \tilde{X}_c$, else $\tilde{C}(\tilde{p}_c, \tilde{t}) = 0$.*

Theorem 20. *The control strategy corresponding to control places for all $\hat{\ell}(\tilde{M}) \in \mathcal{F}$ is maximally per-*

missive, i.e., it does not prevent the unfolding to reach a marking \tilde{M}' with $\hat{\ell}(\tilde{M}') \notin \mathcal{F}$, if the set \mathcal{F} has property reach.

Proof. By construction, each control place for a marking $\hat{\ell}(\tilde{M}) \in \mathcal{F}$ corresponding to configuration \tilde{X} forbids the configuration \tilde{X} and all larger configurations $\tilde{X}' \geq \tilde{X}$. From Theorem 16, the control place only prevents from reaching markings $\hat{\ell}(\tilde{M}) \in \mathcal{F}$. As a result, the control strategy corresponding to control places for all $\hat{\ell}(\tilde{M}) \in \mathcal{F}$ is maximally permissive. \square

We now show that such a controller can be constructed from order 1 unfolding $\tilde{N}_1(M_0)$ including all its cut-off transitions and their output places. The key behind this construction is that avoiding \mathcal{F} can be achieved by one-step look-forward by checking whether firing a new transition leads to a marking in \mathcal{F} .

We first construct the control places that prevent reaching markings in \mathcal{F} in $\tilde{N}_1(M_0)$.

Algorithm 21. Construction of control places for \mathcal{F}

1. Determine a reachable marking \tilde{M} such that $\hat{\ell}(\tilde{M}) \in \mathcal{F}$ and such that no marking \tilde{M}' with $\hat{\ell}(\tilde{M}') \in \mathcal{F}$ and $\text{conf}(\tilde{M}') \subset \text{conf}(\tilde{M})$ exists.
2. If no such marking exists, then stop.
3. Add to $\tilde{N}_1(M_0)$ the control place for \tilde{M} .
4. Goto 1.

The net obtained by adding these control places to the order 1 unfolding is called $\tilde{N}_{1,c}(M_0)$. This net is not necessarily an occurrence net because the control places may contain more than one token. It is however always possible to convert it into an equivalent occurrence net if necessary. This procedure however is not necessary because we can use the net $\tilde{N}_{1,c}(M_0)$ as it is to solve our control problem.

Example 22. Given the net in Fig. 1, assume we want to forbid the set of markings $M \in R(N, M_0)$ such that $M(p_3) + M(p_4) + M(p_7) + M(p_8) = 2$. Clearly $\mathcal{F} = \{\{p_3, p_7\}, \{p_4, p_7\}, \{p_3, p_8\}, \{p_4, p_8\}\}$, and it is not difficult to show that this forbidden set has property REACH for this net.

Consider the order 1 unfolding shown in Figure 3 where to avoid any confusion we have assigned a unique label to each place and transition.

In the following table for each of the forbidden markings M we have shown the corresponding unfolding marking(s) \tilde{M} , the configuration $\text{conf}(\tilde{M})$, the corresponding set of control transitions \tilde{X}_c and finally the control place \tilde{p}_c . A symbol * (resp., **) in the last column denotes a non minimal configuration already forbidden by place \tilde{p}_{c1} (resp., \tilde{p}_{c2}) hence no new place has to be added to the net for preventing it.

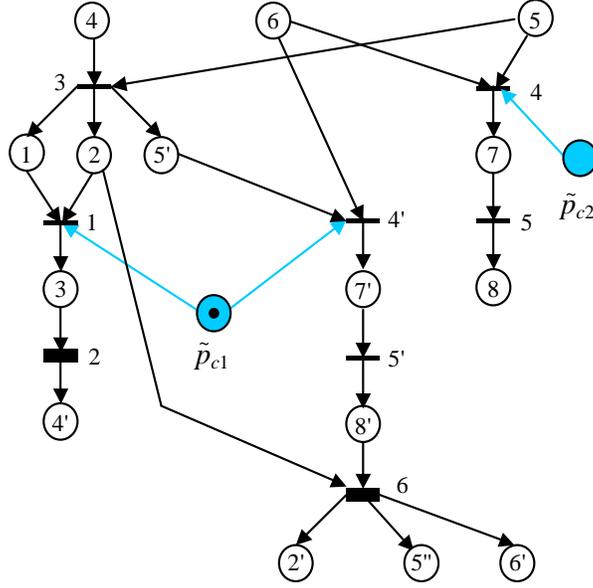


Figure 3: The net $\tilde{N}_{1,c}(M_0)$ in Example 22.

M	\tilde{M}	$conf(\tilde{M})$	\tilde{X}_c	\tilde{p}_c
$\{p_3, p_7\}$	$\{\tilde{p}_3, \tilde{p}'_7\}$	$\{\tilde{t}_3, \tilde{t}_1, \tilde{t}'_4\}$	$\{\tilde{t}_1, \tilde{t}'_4\}$	\tilde{p}_{c1}
$\{p_4, p_7\}$	$\{\tilde{p}_4, \tilde{p}'_7\}$	$\{\tilde{t}_4\}$	$\{\tilde{t}_4\}$	\tilde{p}_{c2}
	$\{\tilde{p}'_4, \tilde{p}'_7\}$	$\{\tilde{t}_3, \tilde{t}_1, \tilde{t}_2, \tilde{t}'_4\}$	$\{\tilde{t}_2, \tilde{t}'_4\}$	*
$\{p_3, p_8\}$	$\{\tilde{p}_3, \tilde{p}'_8\}$	$\{\tilde{t}_3, \tilde{t}_1, \tilde{t}'_4, \tilde{t}'_5\}$	$\{\tilde{t}_1, \tilde{t}'_5\}$	*
$\{p_4, p_8\}$	$\{\tilde{p}_4, \tilde{p}'_8\}$	$\{\tilde{t}_4, \tilde{t}_5\}$	$\{\tilde{t}_5\}$	**
	$\{\tilde{p}'_4, \tilde{p}'_8\}$	$\{\tilde{t}_3, \tilde{t}_1, \tilde{t}_2, \tilde{t}'_4, \tilde{t}'_5\}$	$\{\tilde{t}_2, \tilde{t}'_5\}$	*

The net $\tilde{N}_{1,c}(M_0)$ obtained adding to the order 1 unfolding the control places is shown in Figure 3. Note that place \tilde{p}_{c2} contains no token because its corresponding set of control transitions is a singleton: this means that transition \tilde{t}_4 on tier 1 should never fire. ■

If we want to use the order 1 unfolding for computing a control law for sequences of unbounded length, it is necessary to be able to reset the marking of the unfolding after a sequence containing a cut-off transition has fired. To show how this is possible, we introduce the notion of mirror marking using the concept of *extension of a local configuration* (Esparza *et al.*, 2002) and of equivalent extension.

Definition 23. Let \tilde{M} be a marking of an unfolding net such that $conf(\tilde{M}) = [\tilde{t}] \cup E$ and $[\tilde{t}] \cap E = \emptyset$, i.e. $conf(\tilde{M})$ is an extension of $[\tilde{t}]$ denoted as $[\tilde{t}] \oplus E$. If \tilde{t} is a cut-off transition with \tilde{t}' as its mirror transition in $\tilde{N}_1(M_0)$, we define the mirror marking of \tilde{M} as the marking of the configuration $[\tilde{t}'] \oplus E'$ where E' is the equivalent extension of E for $[\tilde{t}']$.

As an example, consider the unfolding net of Figure 3 before adding \tilde{p}_{c1} and \tilde{p}_{c2} . The configuration of the marking $\tilde{M} = \{\tilde{p}'_4, \tilde{p}'_7\}$ is $conf(\tilde{M}) = \{\tilde{t}_3, \tilde{t}_1, \tilde{t}'_4, \tilde{t}_2\}$ containing one cut-off transition \tilde{t}_2 with a

local configuration $[\tilde{t}_2] = \{\tilde{t}_3, \tilde{t}_1, \tilde{t}_2\}$ and extension $E = \{\tilde{t}'_4\}$. The mirror transition of \tilde{t}_2 is $\tilde{t}' = \varepsilon$, the equivalent extension $E' = \{\tilde{t}_4\}$ and the mirror marking is the marking of configuration $[\varepsilon] \cup \{\tilde{t}_4\} = \{\tilde{t}_4\}$ which is $\{\tilde{p}_4, \tilde{p}_7\}$.

The concept of mirror marking can also be extended to a control place.

Definition 24. *Let \tilde{M} be a marking of $\tilde{N}_{1,c}(M_0)$ such that $\text{conf}(\tilde{M}) = [\tilde{t}] \cup E$ and $[\tilde{t}] \cap E = \emptyset$. If \tilde{t} is a cut-off transition with \tilde{t}' as its mirror transition, the mirror marking of a control place \tilde{p}_c is $\tilde{M}'(\tilde{p}_c) = \tilde{M}_0(\tilde{p}_c) + \tilde{C}(\tilde{p}_c, \cdot)([\tilde{t}'] \oplus E')$, where E' is the equivalent extension of E for $[\tilde{t}']$.*

According to the previous definition, after a cut-off transition fires, the control places should get back all those tokens that have been taken by the firing of $[\tilde{t}] - [\tilde{t}']$.

As an example, consider the control net of Figure 3 and the marking \tilde{M} obtained firing cut-off transition \tilde{t}_2 . Such a marking can be written as: $\tilde{M}(\tilde{P}) = \{\tilde{p}'_4, \tilde{p}'_5, \tilde{p}_6\}$ — considering its projection over the places of the unfolding net (excluding the control places) — while $\tilde{M}(\tilde{p}_{c1}) = \tilde{M}(\tilde{p}_{c2}) = 0$.

The configuration of this marking is $\text{conf}(\tilde{M}) = [\tilde{t}_2] = \{\tilde{t}_3, \tilde{t}_1, \tilde{t}_2\}$ with extension $E = \emptyset$. The mirror transition of \tilde{t}_2 is $\tilde{t}' = \varepsilon$ hence the mirror marking relative to the places of the unfolding is: $\tilde{M}'(\tilde{P}) = \{\tilde{p}_4, \tilde{p}_5, \tilde{p}_6\}$. The mirror marking of the control place \tilde{p}_{c1} can be computed as follows: $\tilde{M}'(\tilde{p}_{c1}) = \tilde{M}_0(\tilde{p}_{c1}) + \tilde{C}(\tilde{p}_{c1}, \cdot)([\varepsilon]) = 1 + 0 = 1$. Analogously, one can compute $\tilde{M}'(\tilde{p}_{c2}) = 0$.

Definition 25. *The control policy for \mathcal{F} uses the net $\tilde{N}_{1,c}(M_0)$ and can be defined as follows.*

1. *The plant and the net $\tilde{N}_{1,c}(M_0)$ are initialized with the respective initial marking.*
2. *Compute a control pattern as follows: if \tilde{T}_e is the set of transitions enabled in $\tilde{N}_{1,c}(M_0)$, the set of transitions that are enabled by the controller on the plant is $T_e = \ell(\tilde{T}_e)$.*
3. *If a transition t fires in the plant, the unique transition $\tilde{t} \in \ell^{-1}(t)$ enabled in $\tilde{N}_{1,c}(M_0)$ is fired. After the firing of \tilde{t} , the marking of the unfolding is set to the related mirror marking if \tilde{t} is a cut-off transition.*
4. *Goto 2.*

As an example, let us consider the control actions determined using the unfolding net of Figure 3 at different steps of the firing sequence $t_3 t_4 t_5 t_6 t_1$. In the following table σ denotes the sequence generated by the plant, $\tilde{M}(\tilde{P})$ denotes the corresponding marking relative to the places of the unfolding (excluding the control places), $\tilde{M}(\tilde{p}_{c1})$ and $\tilde{M}(\tilde{p}_{c2})$ denote the corresponding marking of the control places, $\text{enab}(\sigma)$ denotes the set of transitions enabled in the plant after the firing of σ , while $\phi(\sigma)$ denotes the set of transitions enabled by the controller after the firing of σ .

σ	$\tilde{M}(\tilde{P})$	$\tilde{M}(\tilde{p}_{c1})$	$\tilde{M}(\tilde{p}_{c2})$	$enab(\sigma)$	$\phi(\sigma)$
ε	$\{\tilde{p}_4, \tilde{p}_5, \tilde{p}'_5\}$	1	0	$\{t_3, t_4\}$	$\{t_3\}$
t_3	$\{\tilde{p}_1, \tilde{p}_2, \tilde{p}'_5, \tilde{p}_6\}$	1	0	$\{t_1, t_4\}$	$\{t_1, t_4\}$
t_3t_4	$\{\tilde{p}_1, \tilde{p}_2, \tilde{p}'_7\}$	0	0	$\{t_1, t_5\}$	$\{t_5\}$
$t_3t_4t_5$	$\{\tilde{p}_1, \tilde{p}_2, \tilde{p}'_8\}$	0	0	$\{t_1, t_6\}$	$\{t_6\}$
$t_3t_4t_5t_6$	$\{\tilde{p}_1, \tilde{p}'_2, \tilde{p}''_5, \tilde{p}'_6\}$ \downarrow (mirror)	0	0		
	$\{\tilde{p}_1, \tilde{p}_2, \tilde{p}'_5, \tilde{p}_6\}$	1	0	$\{t_1, t_4\}$	$\{t_1, t_4\}$
$t_3t_4t_5t_6t_1$	$\{\tilde{p}_3, \tilde{p}'_5, \tilde{p}_6\}$	0	0	$\{t_2, t_4\}$	$\{t_2\}$

Theorem 26. *The control policy of Definition 25 is maximally permissive.*

Proof. Similar to the proof of Theorem 20, a marking \tilde{M} of $\tilde{N}_1(M_0)$ is forbidden by control places of Algorithm 21 if and only if $\hat{\ell}(\tilde{M}) \in \mathcal{F}$. Since a marking \tilde{M} obtained by a cut-off transition \tilde{t} is replaced by its mirror marking \tilde{M}' , we need to prove that \tilde{M}' is also permitted by control places. This is true since \tilde{M} is accepted by control places which implies $\hat{\ell}(\tilde{M}) \notin \mathcal{F}$ and $\tilde{M} =_P \tilde{M}'$. The maximal permissiveness is a consequence of the completeness of the unfolding. \square

A final comment to conclude this section. In general Algorithm 21 requires an exhaustive search of all forbidden markings. However, assume that, as in (1), the set of forbidden markings is given by a linear set. Under this assumption the following theorem holds.

Theorem 27. *Given a forbidden set \mathcal{F} of the form given in eq. (1), the set of minimal configurations to be forbidden in Algorithm 21 are the minimal vectors \tilde{X} satisfying the constraint set*

$$\begin{cases} \tilde{M} = \tilde{M}_0 + \tilde{C}_{1,c}\tilde{X} & (a) \\ \tilde{A}\tilde{M} \leq b & (b) \\ \tilde{M} \in \mathbb{N}^{\tilde{m}}; \tilde{X} \in \mathbb{N}^{\tilde{n}}. \end{cases} \quad (2)$$

where $\tilde{C}_{1,c}$ is the incidence matrix of the order 1 unfolding augmented with existing control places.

Proof. The state equations can be used to characterize reachability in (a) thanks to Corollary 12. \square

Example 28. The set of forbidden markings of the previous example

$$\mathcal{F} = \{M \in \mathbb{N}^m \mid M(p_3) + M(p_4) + M(p_7) + M(p_8) = 2\},$$

is in the form given by (1).

We start by solving for the uncontrolled net the IPP

$$\begin{cases} \min \sum_{i \in \tilde{T}} \tilde{X}(\tilde{t}) \\ \tilde{M} = \tilde{M}_0 + \tilde{C}_1\tilde{X} \\ \tilde{M}(\tilde{p}_3) + \tilde{M}(\tilde{p}_4) + \tilde{M}(\tilde{p}'_4) + \tilde{M}(\tilde{p}_7) + \tilde{M}(\tilde{p}'_7) + \tilde{M}(\tilde{p}_8) + \tilde{M}(\tilde{p}'_8) = 2 \\ \tilde{M} \in \mathbb{N}^{\tilde{m}}; \tilde{X} \in \mathbb{N}^{\tilde{n}}. \end{cases}$$

and determine configuration $\tilde{X}_2 = \{t_4\}$ to be forbidden.

We add to the net the corresponding control place $\tilde{p}_{c,2}$ in Figure 3 and for the controlled net thus obtained, whose incidence matrix we denote $\tilde{C}'_{1,c}$, we solve

$$\left\{ \begin{array}{l} \min \sum_{\tilde{t} \in \tilde{T}} \tilde{X}(\tilde{t}) \\ \tilde{M} = \tilde{M}_0 + \tilde{C}'_{1,c} \tilde{X} \\ \tilde{M}(\tilde{p}_3) + \tilde{M}(\tilde{p}_4) + \tilde{M}(\tilde{p}'_4) + \tilde{M}(\tilde{p}_7) + \tilde{M}(\tilde{p}'_7) + \tilde{M}(\tilde{p}_8) + \tilde{M}(\tilde{p}'_8) = 2 \\ \tilde{M} \in \mathbb{N}^{\tilde{m}}; \tilde{X} \in \mathbb{N}^{\tilde{n}}. \end{array} \right.$$

and determine configuration $\tilde{X}_1 = \{t_3, t_1, t'_4\}$ to be forbidden.

We add to the net the corresponding control place $\tilde{p}_{c,1}$ in Figure 3 and for the controlled net thus obtained, whose incidence matrix we denote $\tilde{C}''_{1,c}$, we can check that constraint set (2) is unfeasible, hence no other configuration should be forbidden. \blacksquare

6 Control policy for \mathcal{F}_I

Since \mathcal{F}_I also has the REACH property, the control policy for \mathcal{F} applies if \mathcal{F}_I is known and the order 1 unfolding is enough. Unfortunately, for most control problems, \mathcal{F}_I is not given and has to be determined.

To check whether a reachable marking $\hat{\ell}(\tilde{M}) \in \mathcal{F}_I$, we need to check whether \mathcal{F} is avoidable starting from \tilde{M} . Order 1 unfolding is no longer enough as it does not allow the reachability analysis for all reachable markings. This is possible with order 2 unfolding thanks to the following theorem.

Theorem 29. *Given a net system $\langle N, M_0 \rangle$, let $M \in R(N, M_0)$ be a reachable marking and let $\tilde{M} \in R(\tilde{N}_1(M_0))$ be a marking of the unfolding such that $\hat{\ell}(\tilde{M}) = M$. Then the order 1 unfolding $\tilde{N}_1(M)$ of net N with initial marking M is a subnet of $N_2(M_0)$ starting at \tilde{M} .*

Proof. Consider any configuration \tilde{X} of $\tilde{N}_1(M_0)$ corresponding to marking \tilde{M} . Considering the order 1 unfolding $\tilde{N}_1(\tilde{M})$ starting at \tilde{M} . For any configuration of \tilde{Y} of $\tilde{N}_1(\tilde{M})$, from the completeness of the unfolding net $\tilde{N}(M_0)$, $\tilde{X} + \tilde{Y}$ is a configuration of $\tilde{N}(M_0)$ and $\tilde{N}_1(\tilde{M})$ is a subnet of $\tilde{N}(M_0)$. The theorem is proved if any transition \tilde{t} in \tilde{Y} is either a cut-off transition of $\tilde{N}_2(M_0)$ or its local configuration $[\tilde{t}]$ does not contain any cut-off transition of $\tilde{N}_2(M_0)$. For this purpose assume that there exists a transition \tilde{t} in \tilde{Y} such that its local configuration $[\tilde{t}]$ contains a cut-off transition \tilde{w} of $\tilde{N}_2(M_0)$. Of course \tilde{w} belongs to $\tilde{N}_1(\tilde{M})$ as well. From the definition of $\tilde{N}_2(M_0)$, there exists another transition \tilde{w}' such that (a) either \tilde{w}' does not belong to $\tilde{N}_1(\tilde{M})$ or it is a cut-off transition of $\tilde{N}_1(\tilde{M})$; (b) \tilde{w}' has a smaller configuration: $[\tilde{w}'] \subset [\tilde{w}]$; (c) the markings reached firing the two configurations are equivalent: $\tilde{M}([\tilde{w}']) =_P \tilde{M}([\tilde{w}])$. From the above definition, \tilde{w}' is a transition of the local configuration $[\tilde{t}]$. Further by construction \tilde{t} is not in conflict with any transition in \tilde{X} and hence \tilde{w}' is not in conflict with \tilde{X} . As a result, $\tilde{Z} = \tilde{X} \cup [\tilde{w}']$ is configuration, $\tilde{Z} \subset \tilde{X}$ and \tilde{w}' is a transition of $\tilde{N}_1(\tilde{M})$. Similarly \tilde{w} is a transition of $\tilde{N}_1(\tilde{M})$ and it is a cut-off transition of $\tilde{N}_1(\tilde{M})$. Because \tilde{t} follows \tilde{w} , it cannot be in $\tilde{N}_1(\tilde{M})$. This contradicts the fact that \tilde{t} is a transition of $\tilde{N}_1(\tilde{M})$ and concludes the proof. \square

Hence, if we identify in the order 2 all markings \tilde{M} such that $\hat{\ell}(\tilde{M}) \in \mathcal{F}$, then we can easily identify, by reachability analysis, all markings in \mathcal{F}_I . Structural analysis may also be used for this purpose, as we discuss in the following section.

We conclude this section showing that the set \mathcal{F}_I can be prevented by a non deadlocking supervisor, i.e., the corresponding supervisor is such that the controlled net does not contain controller induced deadlocks.

Theorem 30. *Let $\tilde{N}_{1,c}(M_0)$ be the controlled unfolding net in all markings \tilde{M} such that $\hat{\ell}(\tilde{M}) \in \mathcal{F}_I$ are forbidden by their related control places. Then there exist no dead marking in \tilde{M} of $\tilde{N}_{1,c}(M_0)$ unless it is also a dead marking of $\tilde{N}_1(M_0)$.*

Proof. Let us assume, it is possible to reach in the controlled net a marking \tilde{M} that is a control induced dead marking, i.e., a marking that is dead because of the controller but that is not dead in the order 1 unfolding. Since the control places only forbid transitions firings that lead to \mathcal{F}_I , then without control all transitions enabled at \tilde{M} would lead to a marking in \mathcal{F}_I in one step. By definition, this implies that $\hat{\ell}(\tilde{M}) \in \mathcal{F}_I$. But this is a contradiction, because we have assumed that no markings in \mathcal{F}_I is reachable in the controlled net. \square

7 Deadlock avoidance control

We consider the particular case in which the set of forbidden markings \mathcal{F} is the set of dead markings. Hence the set \mathcal{F}_I is the set of the impending deadlocks and a control law that prevent this set corresponds to a maximally permissive nonblocking supervisor. Under this assumption, we show that there exists an efficient approach based on linear algebra to identify markings in \mathcal{F}_I and to prevent them with control places.

In this section, when we need not distinguish between order 1 and order 2 unfolding we omit the subscript 1 or 2. Thus we denote an unfolding \tilde{N} while its incidence matrix is the $\tilde{m} \times \tilde{n}$ matrix \tilde{C} . Similarly, the controlled net with the addition of \tilde{m}_c control places is denoted \tilde{N}_c while its incidence matrix is the $(\tilde{m} + \tilde{m}_c) \times \tilde{n}$ matrix \tilde{C}_c .

We first observe that the controlled net is an acyclic net. In fact the unfolding is acyclic by construction, and the addition of control places does not modify this property, because each control place has only output arcs. Thus Proposition 11 and Corollary 12 also holds for the controlled net.

Secondly, we discuss how it is possible to give a linear algebraic characterization of the set of deadlock marking. The following result holds.

Proposition 31. *Given an unfolding net $\tilde{N}(\tilde{M}_0)$, we have that a marking \tilde{M} is dead if and only if for all $\tilde{t} \in \tilde{T}$ if holds*

$$\sum_{\tilde{p} \in \bullet \tilde{t}} \tilde{M}(\tilde{p}) \leq |\bullet \tilde{t}| - 1.$$

Proof. The result follows from the fact that the unfolding is a safe net, hence \tilde{t} if enabled if and only if $\tilde{M}(\tilde{p}) = 1$ for all $\tilde{p} \in \bullet \tilde{t}$. \square

This result does not hold for the controlled net, because the control places are not necessarily safe. However, the following results holds.

Proposition 32. *Given a controlled net $\tilde{N}_c(\tilde{M}_{c,0})$, let \tilde{P} be the set of places of the unfolding net, and \tilde{P}_c the set of control places. Given any marking \tilde{M} , we can associate to each place $\tilde{p}_c \in \tilde{P}_c$ a binary counter $\mu(\tilde{p}_c) \in \{0, 1\}$ that satisfies the following equations:*

$$\mu(\tilde{p}_c) \leq \tilde{M}(\tilde{p}_c) \quad \text{and} \quad \tilde{M}_{c,0}(\tilde{p}_c)\mu(\tilde{p}_c) \geq \tilde{M}(\tilde{p}_c). \quad (3)$$

Then a marking \tilde{M} is dead if and only if for all $\tilde{t} \in \tilde{T}$ it holds

$$\sum_{\tilde{p} \in \tilde{P} \cap \bullet \tilde{t}} \tilde{M}(\tilde{p}) + \sum_{\tilde{p}_c \in \tilde{P}_c \cap \bullet \tilde{t}} \mu(\tilde{p}_c) \leq |\bullet \tilde{t}| - 1.$$

Proof. We first observe that the first equation (3) implies that $\mu(\tilde{p}) = 0$ if $\tilde{M}(\tilde{p}_c) = 0$, while the second equation (3) implies that $\mu(\tilde{p}) = 1$ if $\tilde{M}(\tilde{p}_c) > 0$ (note that by construction the control place is such that $\tilde{M}_{c,0}(\tilde{p}_c) \geq \tilde{M}(\tilde{p}_c)$), i.e., $\mu(\tilde{p}) = 1$ if and only if \tilde{p} is marked. The results follows because \tilde{t} is enabled if and only if all its input places are marked. \square

Our third and final preliminary result characterizes redundant control places, i.e., places that can be removed without changing the behavior of the controlled net.

Definition 33. *Given a controlled net $\tilde{N}_c(\tilde{M}_{c,0})$, let $\tilde{N}'_c(\tilde{M}'_{c,0})$ be the net obtained from \tilde{N}_c removing a control place \tilde{p}'_c . We say that place \tilde{p}'_c is redundant in $\tilde{N}_c(\tilde{M}_{c,0})$ if for all reachable markings $\tilde{M} \in R(\tilde{N}_c(\tilde{M}_{c,0}))$ and for all transitions $\tilde{t} \in \tilde{p}' \bullet$ it holds $(\forall \tilde{p}_c \in \bullet \tilde{t} \setminus \{\tilde{p}'_c\}) \tilde{M}(\tilde{p}_c) > 0 \implies \tilde{M}(\tilde{p}'_c) > 0$. \blacksquare*

Proposition 34. *With the notation of the previous definition, place \tilde{p}'_c is redundant in $\tilde{N}_c(\tilde{M}_{c,0})$ if and only if the following integer programming problem (IPP)*

$$\begin{cases} \min J = C(\tilde{p}'_c, \cdot) \tilde{X} \\ \text{s.t.} \quad \tilde{M}'_{c,0} + \tilde{C}' \tilde{X} \geq 0 \end{cases}$$

— where \tilde{C} and \tilde{C}' are, respectively, the incidence matrices of \tilde{N}_c and \tilde{N}'_c — has optimal solution J^* such that $\tilde{M}_{c,0}(\tilde{p}') + J^* \geq 0$.

Proof. By Proposition 11, any vector \tilde{X} satisfying the IPP corresponds to a firable sequence of $\tilde{N}'_c(\tilde{M}'_{c,0})$. This sequence is never disabled by place \tilde{p}'_c in $\tilde{N}_c(\tilde{M}_{c,0})$ if $\tilde{M}_{c,0}(\tilde{p}'_c) + J^* \geq 0$. Hence $\tilde{N}'_c(\tilde{M}'_{c,0})$ and $\tilde{N}_c(\tilde{M}_{c,0})$ admit the same set of firable sequences. \square

We can finally outline an algorithm to design a maximally permissive deadlock avoidance controller for a given safe net system $\langle N, M_0 \rangle$.

Algorithm 35. Deadlock avoidance controller

1. Construct the order 2 unfolding $\tilde{N}_2(\tilde{M}_0)$.
2. Determine the set of dead markings of $\tilde{N}_2(\tilde{M}_0)$, excluding the markings that include the output places \tilde{P}_{out} of the cut-off transitions of the order 2 unfolding³. This set corresponds to the feasible

³These markings even if they are dead in $\tilde{N}_2(\tilde{M}_0)$, do not necessarily correspond to dead markings in the original net.

solutions \tilde{M} of the following constraint set

$$\left\{ \begin{array}{l} \tilde{M} = \tilde{M}_0 + \tilde{C}_2 \tilde{X} \\ \sum_{\tilde{p} \in \bullet \tilde{t}} \tilde{M}(\tilde{p}) \leq |\bullet \tilde{t}| - 1 \quad (\forall \tilde{t} \in \tilde{T}) \\ \tilde{M}(\tilde{p}) = 0 \quad (\forall \tilde{p} \in \tilde{P}_{out}) \\ \tilde{M} \in \mathbb{N}^{\tilde{m}}, \tilde{X} \in \mathbb{N}^{\tilde{n}} \end{array} \right.$$

and for each marking \tilde{M} add to the unfolding the corresponding set of control places to obtain a net $\tilde{N}_{2,c}(\tilde{M}_{c,0})$.

3. Determine the set of dead markings of $\tilde{N}_{2,c}(\tilde{M}_{c,0})$ as the set of feasible solutions \tilde{M} of the following constraint set

$$\left\{ \begin{array}{l} \tilde{M} = \tilde{M}_{c,0} + \tilde{C}_{2,c} \tilde{X} \\ \mu(\tilde{p}_c) \leq \tilde{M}(\tilde{p}_c) \quad (\forall \tilde{p}_c \in \tilde{P}_c) \\ \tilde{M}_{c,0}(\tilde{p}_c) \mu(\tilde{p}_c) \geq \tilde{M}(\tilde{p}_c) \quad (\forall \tilde{p}_c \in \tilde{P}_c) \\ \sum_{\tilde{p} \in \tilde{P} \cap \bullet \tilde{t}} \tilde{M}(\tilde{p}) + \sum_{\tilde{p}_c \in \tilde{P}_c \cap \bullet \tilde{t}} \mu(\tilde{p}_c) \leq |\bullet \tilde{t}| - 1 \quad (\forall \tilde{t} \in \tilde{T}) \\ \tilde{M}(\tilde{p}) = 0 \quad (\forall \tilde{p} \in \tilde{P}_{out}) \\ \tilde{M} \in \mathbb{N}^{\tilde{m}}, \tilde{X} \in \mathbb{N}^{\tilde{n}}, \mu \in \{0, 1\}^{\tilde{m}_c} \end{array} \right.$$

4. If the set of dead markings determined at the previous step is not empty, add to $\tilde{N}_{2,c}$ the corresponding control places and go to 3.
5. Let $\tilde{N}_{1,c}$ be the net obtained from $\tilde{N}_{2,c}$ removing all places and transitions that do not belong to the order 1 unfolding, and removing all control places that have arcs going to a transition that has been removed.
6. Check all control places of $\tilde{N}_{1,c}$ for redundancy, using the IPP of Proposition 34, and remove the redundant ones. ■

It is not difficult to show that the control net computed by the previous algorithm corresponds to a maximally permissive deadlock avoidance controller for a given safe net system $\langle N, M_0 \rangle$. We give here a sketch of this proof.

- First, let us observe that all markings forbidden in step 2 are dead markings of the original net, and let $\tilde{\mathcal{D}}$ be this set of markings. The k -th iteration of step 3 forbids all markings of $\tilde{N}_2(\tilde{M}_0)$ that are k steps from markings in $\tilde{\mathcal{D}}$, hence by repeating step 3 one forbids all markings that unavoidably leads to markings in $\tilde{\mathcal{D}}$ in a finite number of steps. Let $\tilde{\mathcal{D}}^*$ be the set of markings that are forbidden either in step 2 or in step 3: clearly these markings are dead markings or impending dead markings that should be forbidden by a deadlock avoidance controller.
- On the other hand, according to Theorem 29, any order 1 unfolding marking \tilde{M} has its order 1 unfolding net as subnet of $\tilde{N}_2(\tilde{M}_0)$. As a result, any order 1 unfolding marking \tilde{M} corresponding to a dead marking or to an impending dead marking of $\langle N, M_0 \rangle$ belongs to the set $\tilde{\mathcal{D}}^*$ and is hence forbidden by the controller of Algorithm 35.

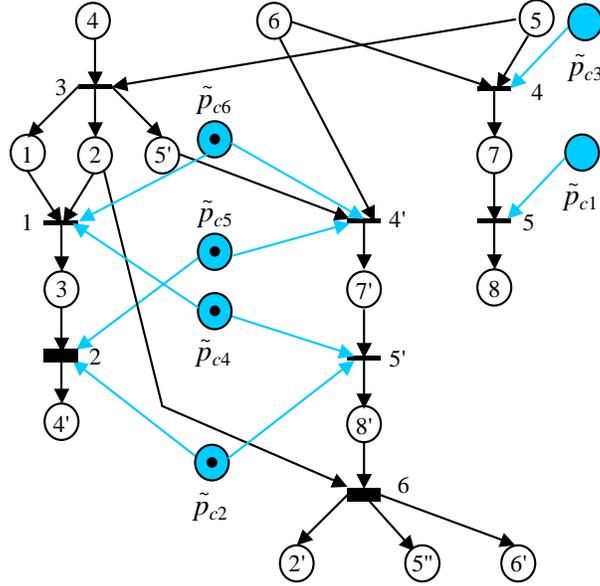


Figure 4: The net $\tilde{N}_{1,c}$ in Example 36 before removing the redundant control places.

- The above two results show that any order 1 unfolding transition enabled by the controller designed with the previous algorithm does not leads to a marking in $\tilde{\mathcal{D}}^*$. Furthermore, it is clear that a mirror marking following of an order 1 cut-off transition belongs to $\tilde{\mathcal{D}}^*$ iff its original marking belongs to $\tilde{\mathcal{D}}^*$. These two results show the correctness of the controller designed with Algorithm 35, while the completeness of the order 1 unfolding ensures the maximum permissiveness of the controller.

The net constructed with the previous algorithm can be used to compute a maximally permissive non-blocking control policy, as explained in Definition 25.

Example 36. The net in Figure 1 is blocking. Using the previous algorithm we first construct its order 2 unfolding (see Figure 2) and at step 2 identify three blocking markings: $\{\tilde{p}_4, \tilde{p}_8\}$, $\{\tilde{p}'_4, \tilde{p}'_8\}$ and $\{\tilde{p}''_4, \tilde{p}''_8\}$. All correspond to the unique dead marking of the original net that marks places p_4 and p_8 . Iterating at step 3 of the algorithm we identify new control induced markings, and add the corresponding control places to the order 2 unfolding. At step 5, removing the second order subnet, we obtain the net $\tilde{N}_{1,c}$ shown in Figure 4. Finally at step 5 we eliminate the redundant places: only places $\tilde{p}_{c,3}$ and $\tilde{p}_{c,6}$ remain at the end of the algorithm. ■

8 Conclusions

In this paper we have used the technique of unfolding to design maximally permissive supervisors for safe Petri nets assuming that the specification is given by a set of forbidden markings \mathcal{F} with property REACH. The control structure we have derived takes the form of a set of places to be added to the unfolding. The unfolding net with the addition of control places is used to compute a maximally permissive control policy.

The approach has also been extended to the problem of preventing a larger set \mathcal{F}_I of impending forbidden

marking. This is a superset of the forbidden markings that also includes all those markings from which - unless the supervisor blocks the plant - a marking in \mathcal{F} is inevitably reached in a finite number of steps.

Although in the general case it may be necessary to enumerate all forbidden markings, the technique can be efficiently applied to the design of maximally permissive supervisors for deadlock avoidance by means of structural analysis.

There are some lines for future research that are still open.

- In many cases it may be possible to find equivalent control structure to be added to the original net rather than to the unfolding.
- The approach may also be extended to nets with uncontrollable transitions, that cannot be disabled by a supervisor.

References

- Aghasaryan, A., E. Fabre, A. Benveniste, R. Boubour and C. Jard (1998). Fault detection and diagnosis in distributed systems: an approach by partially stochastic Petri nets. *Discrete Events Dynamical Systems* **8**, 203–231.
- Benveniste, A., E. Fabre and S. Haar (2003a). Markov nets: probabilistic models for distributed and concurrent systems. *IEEE Trans. on Automatic Control* **48**(11), 1936–1950.
- Benveniste, A., E. Fabre, S. Haar and C. Jard (2003b). Diagnosis of asynchronous discrete event systems, a net unfolding approach. *IEEE Trans. on Automatic Control* **48**(5), 714–727.
- Chu, F. and X. Xie (1997). Deadlock analysis of Petri nets using siphons and mathematical programming. *IEEE Trans. on Automation* **13**(6), 793–804.
- Esparza, J., S. Römer and W. Vogler (2002). An improvement of mcmillan’s unfolding algorithm. *Formal Methods in System Design* **20**, 285–310.
- Ezpeleta, J., J.M. Colom and J. Martinez (1995). A Petri net based deadlock prevention policy for flexible manufacturing systems. *IEEE Trans. on Automation* **11**(3), 173–184.
- Giua, A. and X. Xie (2004). Control of safe ordinary Petri nets with marking specifications using unfolding. In: *Proc. IFAC WODES’04: 7th Work. on Discrete Event Systems*. Reims, France. pp. 61–66.
- Giua, A. and X. Xie (2005). Nonblocking control of Petri nets using unfolding. In: *Proc. 16th IFAC World Congress*. Prague, Czech Republic.
- Giua, A., F. DiCesare and M. Silva (1992). Generalized mutual exclusion constraints on nets with uncontrollable transitions. In: *Proc. 1992 IEEE Int. Conf. on Systems, Man, and Cybernetics*. pp. 974–979.
- Godefroid, P. (1996). *Partial-order methods for the verification of concurrent systems - an approach to the state-explosion problem*. Vol. 1032 of *Lecture Notes in Computer Science*. Springer Verlag.

- He, K.X and M.D. Lemmon (2000). Liveness verification of discrete-event systems modeled by n-safe ordinary Petri nets. In: *21st Int. Conf. on Application and Theory of Petri Nets (ICATPN 2000)*, Aarhus, Denmark. Vol. 1825 of *Lecture Notes in Computer Science*. Springer Verlag. pp. 227–243.
- He, K.X. and M.D. Lemmon (2002). Liveness-enforcing supervision of bounded ordinary Petri nets using partial order methods. *IEEE Trans. on Automatic Control* **47**(7), 1042–1055.
- Hellgren, A., M. Fabian and B. Lennartson (1999). Deadlock detection and controller synthesis for production systems using partial order techniques. In: *Proc. Conference on Control Applications*. Kohala Coast, Hawaii, USA. pp. 1472–1477.
- McMillan, K.L. (1995). A technique of state space search based on unfolding. *Formal Methods in System Design* **6**(1), 45–65.
- Murata, T. (1989). Petri nets: Properties, analysis and applications. *Proceedings of the IEEE* **77**(4), 541–580.
- Neumair, C. (2002). Finite unfoldings of unbounded Petri nets. *Petri Net Newsletter* (63), 5–10.
- Park, J. and S.A. Reveliotis (2001). Deadlock avoidance in sequential resource allocation systems with multiple resource acquisitions and flexible routings. *IEEE Trans. on Automatic Control* **46**(10), 1572–1583.
- Ramadge, P. and W.M. Wonham (1989). Control of discrete event system. *Proceedings of the IEEE* **77**(1), 81–98.
- Valmari, A. (1991). Stubborn sets for reduced state space generation. In: *Advances in Petri Nets 1990*. Vol. 483 of *Lecture Notes in Computer Science*. Springer-Verlag. pp. 491–515.
- Valmari, A. (1994). State of the art report: Stubborn sets. *Petri Net Newsletter* (46), 6–14.
- Xie, X. and A. Giua (2004). Counterexamples to «Liveness-enforcing supervision of bounded ordinary Petri nets using partial order methods». *IEEE Trans. on Automatic Control* **49**(7), 1217–1219.