

Observability of Place/Transition Nets

Alessandro Giua, Carla Seatzu

Department of Electrical and Electronic Engineering,

University of Cagliari, Piazza d'Armi — 09123 Cagliari, Italy

Tel: +39 (70) 675-5892. Fax: +39 (70) 675-5900. Email: {giua,seatzu}@diee.unica.it.

Abstract

In this paper we discuss the problem of estimating the marking of a Place/Transition net based on event observation. We assume that the net structure is known while the initial marking is totally or partially unknown.

We give algorithms to compute a marking estimate that is a lower bound of the actual marking. The special structure of Petri nets allows us to use a simple linear algebraic formalism for estimate and error computation.

The error between actual marking and estimate is a monotonically non-increasing function of the observed word length, and words that lead to null error are said complete. We define several observability properties related to the existence of complete words, and show how they can be proved. To prove some of them we also introduce a useful tool, the observer coverability graph, i.e., the usual coverability graph of a Place/Transition net augmented with a vector that keeps track of the estimation error on each place of the net.

Finally, we show how the estimate generated by the observer may be used to design a state feedback controller for forbidden marking specifications.

Published as: A. Giua, C. Seatzu, "Observability of Place/Transition Nets," *IEEE Transactions on Automatic Control*, Vol. 47, No. 9, pp. 1424 -1437, September 2002.

1 Introduction

This paper deals with the problem of estimating the marking of a Place/Transition (P/T) net based on the observation of transition firings and presents a set of analytical tools to determine several observability properties. An observer constructed following this approach can also be used in a state feedback control loop, as discussed in the final part of the paper.

This framework provides a useful paradigm that can be applied to different settings, from discrete event control, to failure diagnosis and error recovery. The assumption that only event occurrences, i.e., transition firings, may be observed — while the plant state, i.e., the marking, cannot — is common in discrete event control. The assumption that the state of the plant is not known (or is only partially known) is natural during error recovery. Consider for instance the case of a plant remotely controlled: if the communication fails the state may evolve and when the communication is re-established the state will be at best partially known. In a manufacturing environment, one may consider the case in which resources (i.e., tokens) enter unobserved, or in which we know how many resources have entered the system but not their exact location.

When the structure and the initial marking of a P/T net is known, the knowledge of the transition firings is sufficient to reconstruct the marking that each new firing yields. In this work we assume that only the net structure is known and consider two cases: (a) the initial marking is not known; (b) the initial marking is known to belong to a “macromarking”, i.e., we know the token contents of subsets of places but not the exact token distribution. Case (b) can in effect be seen as a generalization of case (a) but we preferred to handle the two cases separately.

In both cases we show how it is possible to estimate the actual marking of the net based on the observation of a word of events (i.e., transition firings) and we give algorithms for computing the estimates and, in case (b), bounds on the error. The estimate is always a lower bound of the actual marking. The system that computes the estimate is called an observer.

The special structure of Petri nets allows us to use a simple linear algebraic formalism for estimate and error computation. In particular, the set of markings consistent with an observed word, i.e., the set of markings in which the system may actually be given the observed word,

can be easily characterized as a convex set of integers.

The error function between the actual marking and the estimate can be shown to be a monotonically non-increasing function of the observed word length. Observed words that lead to a null error are said to be “complete”. Complete observers are the discrete-event counterpart of asymptotic observers for time-driven systems.

In this paper we define several observability properties and show that they are decidable. In particular we consider two main properties. *Marking observability* (MO) means that there exists at least one word that is complete, while *strong marking observability* (SMO) means that all words can be completed in a finite number of steps into a complete word.

We set up a hierarchy considering the possibility that the two properties are satisfied by a net N starting from an initial marking M_0 , by a net N starting from any marking M reachable from an initial marking M_0 (uniform observability) or by a net N starting from any marking in \mathbb{N}^m (structural observability) where m is the number of places of the net.

To prove some of these properties we introduce a useful tool, the *observer coverability graph*, i.e., the usual coverability graph of a Place/Transition net augmented with a vector that keeps track of the estimation error on each place of the net.

All the considered properties can be proven either by the use of the observer coverability graph, or by reducing them to other decision problems (e.g., home-space properties, marking reachability, existence of repetitive sequences) that can be checked using algorithms well known in the literature and that will not be further discussed in this paper. Although more efficient algorithms may exist to prove all observability properties, our main aim here is to characterize these properties and to show that they are decidable.

Finally, we show how the estimate generated by the observer may be used to design a state feedback controller, that ensures that the controlled system never enters a set of forbidden states. In the discussion we assume that all events are controllable and we focus on a special class of specifications that limit the weighted sum of markings in subset of places. Clearly, the use of marking estimates (as opposed to the exact knowledge of the actual marking of the plant) leads to a worse performance of the closed-loop system in the sense that to rule out the possibility that the plant enter a forbidden marking, the controller may prevent the firing of

transitions whose firing is perfectly legal given the actual marking of the plant.

1.1 Relevant literature

Observability is a fundamental property that has received a lot of attention in the framework of time-driven systems, given the importance of reconstructing plant states that cannot be measured. Although less popular in the case of discrete-event systems, the issue of state estimation and of control under partial state observation has been discussed in the literature.

For systems represented as finite automata, Ramadge [21] was the first to show how an observer could be designed for a partially observed system.

Caines *et al.* [3] showed how it is possible to use the information contained in the past sequence of observations (given as a sequence of observation states and control inputs) to compute the set of consistent states. In [4] the observer output is used to steer the state of the plant to a desired terminal state. The approach of Caines is based on the construction of an observer tree to determine the set of markings consistent with the observed behaviour: the tree contains all consistent markings. A similar approach was also used by Kumar *et al.* [11] when defining observer based dynamic controllers in the framework of supervisory predicate control problems.

Özveren and Willsky [18] propose an approach for building observers that allows one to reconstruct the state of finite automata after a word of bounded length has been observed, showing that an observer may have an exponential number of states.

The main drawback of the automata based approach is the requirement that the set of consistent markings must explicitly be enumerated. It may be useful to pinpoint that on the contrary the procedure we present in this paper simply produces an estimate of the state, while the special structure of a Petri net permits to determine, using linear algebraic tools, if a given marking is consistent without the explicit enumeration of the (possibly infinite) consistent set.

Finally, as an example in which state estimators have been used for fault diagnosis in systems represented as finite state machines, we mention the work of Wang and Schwartz [25]. The purpose of the observer was in this framework not that of reconstructing the state of the

system, but rather that of detecting if the system is faulty and recognizing the fault type.

Very few works dealt with observability in Petri nets. As far as we know, the first one were [7, 9] where preliminary concepts discussed in this paper have been introduced.

Meda and Ramírez [15] used Interpreted Petri nets to model the system and the observer. The main idea is to start the observer with a marking bigger than the real one and then eliminate some tokens until the observer and system markings are equal. Interpreted Petri nets have also been used by Ramirez–Treviño *et al.* in [22] where it was shown that observability defined as in [15] is equivalent to observability in [9] and it was shown how to construct an observer for binary Interpreted Petri nets when the observability property is verified.

The issue of controlling a plant with incomplete (state or event) measurements has also been discussed in the discrete event control literature.

Zhang and Holloway [26] used a Controlled Petri Net model for forbidden state avoidance under partial *event* observation with the assumption that the initial marking be known.

The use of state-feedback control under partial *state* observation has been discussed by Li and Wonham [13, 14] and by Takai *et al.* [23]. In the work of these authors the partial observation is due to a static mask, that maps the plant state space into an observation space. The main focus was in finding necessary and sufficient conditions for the existence of “optimal” state feedback control laws given a mask (optimal means that the resulting closed-loop behavior is the same for the controller with mask and the controller with complete state observation).

Unlike the above approach, the setting we deal with in this paper assumes that the mask is induced by the computed estimate, and it changes as the plant evolves. Initially, when the estimate is crude, it is often the case that these restrictive “optimal” conditions are not verified. We propose a control scheme that tries to make the best use of the available estimate to ensure the correct behaviour of the plant under control.

The notion of macromarking we use in this paper is similar to the notion of *uncertain marking*, as described by Cardoso *et al.* [5]. These authors considered high-level nets as models of manufacturing systems and they assumed as theoretical basis of uncertain markings possibilistic logic. In our approach, on the contrary, we restrict our attention to purely logical

models (P/T nets) and we assume that all possible markings have equal probability.

2 Background

In this section we provide some basic definitions that will be used in the following of the paper. We first recall some basic terminology on Petri nets, then we provide the definition of both linear and semi-linear sets and we recall the main results on decidability of home-space property.

2.1 Petri nets

In this subsection we recall the Petri net formalism used in this paper. For a more comprehensive introduction to Petri nets see [17]. A *Place/Transition net* (P/T net) is a structure $N = (P, T, Pre, Post)$, where P is a set of m places; T is a set of n transitions; $Pre : P \times T \rightarrow \mathbb{N}$ and $Post : P \times T \rightarrow \mathbb{N}$ are the *pre-* and *post-incidence functions* that specify the arcs. The *incidence matrix* of the net is defined as $C(p, t) = Post(p, t) - Pre(p, t)$.

We define $p^\bullet = \{t \in T \mid Pre(p, t) > 0\}$ as the set of output transitions of place p .

A *marking* is a vector $M : P \rightarrow \mathbb{N}$ that assigns to each place of a P/T net a non-negative number of tokens, represented by black dots. 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 if $M \geq Pre(\cdot, t)$ and may fire yielding the marking $M' = M + C(\cdot, t)$. We write $M [w \rangle M'$ to denote that the enabled sequence of transitions w may fire at M yielding M' , or equivalently we use the notation $M' = w(M)$ and $M = w^{-1}(M')$. Moreover, we denote $w(M_0) = M_w$. Finally, we denote as w_0 the sequence of null length. The set of all sequences fireable in $\langle N, M_0 \rangle$ is denoted $L(N, M_0)$ (this is also called the prefix-closed free language of the net). If the firing sequence w is enabled at M_0 , we also say that w is a word in $L(N, M_0)$.

Let $w = t_{\alpha_1}, t_{\alpha_2}, \dots, t_{\alpha_k}$ be a sequence in $L(N, M_0)$. The sequence $w_i = t_{\alpha_1}, \dots, t_{\alpha_i}$ with $i \in \mathbb{N}$ and $i \leq k$ is a prefix of w of length i and we write $w_i \preceq w$.

A marking M is *reachable* in $\langle N, M_0 \rangle$ iff there exists a firing sequence w such that $M_0 [w \rangle M$.

The set of all markings reachable from M_0 defines the reachability set of $\langle N, M_0 \rangle$ and is denoted $R(N, M_0)$.

A *repetitive* sequence w is such that $M[w]M'$ with $M' \geq M$. Then $\forall i \geq 1$, w^i is enabled at M . A repetitive sequence w is said to be *non-stationary* if $M[w]M'$ with $M' \succneq M$: such a sequence strictly increases the token count of one or more places.

Three useful elementary facts about Petri nets that will be used in the paper are the following.

Fact 1. (i) $M \leq M' \implies L(N, M) \subseteq L(N, M')$.

(ii) If w is enabled at M and M' then: $M - M' = w(M) - w(M')$.

(iii) The reachability set $R(N, M_0)$ is infinite iff there exists a non-stationary repetitive sequence in $L(N, M_0)$.

Finally, we denote $\vec{0}_m$ ($\vec{1}_m$) a $m \times 1$ vector of zeros (ones).

2.2 Home space property

Linear and semi-linear sets were firstly introduced in [19] in order to study some problems from formal language theory.

Definition 2. We say that $\mathcal{E} \subseteq \mathbb{N}^m$ is a linear set if there exists some $V \in \mathbb{N}^m$ and a finite set $\{V_1, \dots, V_n\} \subseteq \mathbb{N}^m$ such that

$$\mathcal{E} = \{V' \in \mathbb{N}^m \mid V' = V + \sum_{i=1}^n k_i V_i \text{ with } k_i \in \mathbb{N}\},$$

V is called the base of \mathcal{E} , and V_1, \dots, V_n are called its periods.

A semi-linear set is the finite union of a family of linear sets.

A first result regarding decidability is the following.

Theorem 3 ([6]). Given a net system $\langle N, M_0 \rangle$ and a semi-linear set \mathcal{E} it is decidable if $R(N, M_0) \cap \mathcal{E} = \emptyset$.

Finally, we introduce the definition of home space [16] and an important theorem that will be used when proving some properties of estimates.

Definition 4 ([16]). Let \mathcal{HS} be a set of markings. We say that \mathcal{HS} is a home space of a P/T net $\langle N, M_0 \rangle$ iff $\forall M \in R(N, M_0), \exists M' \in \mathcal{HS}$ such that $M' \in R(N, M)$. If \mathcal{HS} is a singleton,

we call its unique element a home state.

Thus a set of markings \mathcal{HS} is a home space if from any reachable marking it is possible to reach some marking in \mathcal{HS} . Furthermore, there exist special classes of sets for which the home state property is decidable.

Theorem 5 ([6]). *The property of being a home space for finite unions of linear sets having the same periods is decidable.*

3 Marking estimation

In this section we present an algorithm for estimating the state of a net system $\langle N, M_0 \rangle$ whose marking cannot be directly observed under the following assumptions.

- A1) The structure of the net $N = (P, T, Pre, Post)$ is known, while the initial marking M_0 is not.
- A2) The event occurrences (i.e., the transition firings) can be observed.

After the word w has been observed we define the set $\mathcal{M}(w)$ of w consistent markings as the set of all markings in which the system may be given the observed behaviour.

Definition 6. *Given an observed word w , the set of w consistent markings is $\mathcal{M}(w) = \{M \mid \exists M' \in \mathbb{N}^m, M'[w]M\}$.*

Given an evolution of the net $M_{w_0}[t_{\alpha_1}]M_{w_1}[t_{\alpha_2}]\cdots$, we use the following algorithm to compute the estimate μ_{w_i} of each actual marking M_{w_i} based on the observation of the word of events $w_i = t_{\alpha_1}t_{\alpha_2}\cdots t_{\alpha_i}$.

Algorithm 7 (Marking Estimation with Event Observation).

1. Let the initial estimate be $\mu_{w_0} = \vec{0}_m$.
2. Let $i = 1$.
3. Wait until t_{α_i} fires.
4. Update the estimate $\mu_{w_{i-1}}$ to μ'_{w_i} with

$$\mu'_{w_i}(p) = \max\{\mu_{w_{i-1}}(p), Pre(p, t_{\alpha_i})\}.$$

5. Let $\mu_{w_i} = \mu'_{w_i} + C(\cdot, t_{\alpha_i})$.

6. Let $i = i + 1$.

7. Goto 3. ■

Note that in step 4. of the algorithm we update the previously computed estimate $\mu_{w_{i-1}}$, since the firing of t_{α_i} implies that $M_{w_{i-1}} \geq \text{Pre}(\cdot, t_{\alpha_i})$. In the following we will always denote the estimate computed by this algorithm after having observed the word w as μ_w .

The estimate computed by Algorithm 7 is a lower bound of the actual marking of the net.

Proposition 8. *Let $w = t_{\alpha_1} t_{\alpha_2} \cdots \in L(N, M_0)$ be an observed string and w_i its prefix of length i . Then*

$$\forall i, \quad \mu_{w_i} \leq \mu'_{w_{i+1}} \leq M_{w_i}.$$

Proof. Clearly $\mu_{w_i} \leq \mu'_{w_{i+1}}$ for all i . Also t_{α_1} is enabled at $M_0 \equiv M_{w_0}$, hence $M_{w_0} \geq \text{Pre}(\cdot, t_{\alpha_1}) = \mu'_{w_1}$.

By induction, assume $\mu_{w_{i-1}} \leq \mu'_{w_i} \leq M_{w_{i-1}}$. Then $\mu_{w_i} = \mu'_{w_i} + C(\cdot, t_{\alpha_i}) \leq M_{w_{i-1}} + C(\cdot, t_{\alpha_i}) = M_{w_i}$. Finally, $t_{\alpha_{i+1}}$ is enabled at M_{w_i} , hence $M_{w_i} \geq \text{Pre}(\cdot, t_{\alpha_{i+1}})$. This implies that $M_{w_i}(p) - \mu'_{w_{i+1}}(p) = M_{w_i}(p) - \mu_{w_i}(p) \geq 0$ if $\mu'_{w_{i+1}}(p) = \mu_{w_i}(p)$ while $M_{w_i}(p) - \mu'_{w_{i+1}}(p) = M_{w_i}(p) - \text{Pre}(p, t_{\alpha_{i+1}}) \geq 0$ if $\mu'_{w_{i+1}}(p) > \mu_{w_i}(p)$. □

It is possible to give an easy characterization of the set of consistent markings in terms of estimate. Let us first consider the following lemma that states that the minimal initial marking enabling a sequence w on a net N is $w^{-1}(\mu_w)$.

Lemma 9. *Let $w = t_{\alpha_1} t_{\alpha_2} \cdots t_{\alpha_k}$ be a sequence of transitions of a net N . Then $w \in L(N, M_0)$ if and only if $M_0 \geq w^{-1}(\mu_w) \equiv \mu_w - \sum_{i=1}^k C(\cdot, t_{\alpha_i})$.*

Proof. (if) Let $\sigma_i = t_{\alpha_i} t_{\alpha_{i+1}} \cdots t_{\alpha_k}$ be the suffix of w of length $k + 1 - i$. We will show that for all i , $\sigma_i^{-1}(\mu_w) \equiv \mu_w - \sum_{j=i}^k C(\cdot, t_{\alpha_j}) \geq \text{Pre}(\cdot, t_{\alpha_i})$. This implies that $w \in L(N, w^{-1}(\mu_w))$, hence (by fact 1.i) $\forall M_0 \geq w^{-1}(\mu_w), w \in L(N, M_0)$. Clearly, $\sigma_k^{-1}(\mu_w) = \mu_w - C(\cdot, t_{\alpha_k}) = \mu'_{w_k} \geq \text{Pre}(\cdot, t_{\alpha_k})$. By induction, assume now that $\sigma_{i+1}^{-1}(\mu_w) \geq \mu'_{w_i} \geq \text{Pre}(\cdot, t_{\alpha_{i+1}})$. Then, $\sigma_i^{-1}(\mu_w) = \sigma_{i+1}^{-1}(\mu_w) - C(\cdot, t_{\alpha_i}) \geq \mu'_{w_{i+1}} - C(\cdot, t_{\alpha_i}) \geq \mu_{w_i} - C(\cdot, t_{\alpha_i}) = \mu'_{w_i} \geq \text{Pre}(\cdot, t_{\alpha_i})$.

(only if) By contradiction. Let $w \in L(N, M_0)$ with $M_0(p) < w^{-1}(\mu_w)(p)$. Then (by fact 1.ii) $w(M_0)(p) < \mu_w(p)$ and this violates Proposition 8. □

This lemma leads to the following theorem.

Theorem 10. *Given an observed word $w \in L(N, M_0)$ and the corresponding estimated marking μ_w computed by Algorithm 7, the set of w consistent markings is*

$$\mathcal{M}(w) = \{M \in \mathbb{N}^m \mid M \geq \mu_w\}.$$

Proof. $w \in L(N, M_0) \iff$ (by Lemma 9) $M_0 \geq w^{-1}(\mu_w) \iff$ (by fact 1.ii) $M_w \geq \mu_w$. \square

It is also possible to define a meaningful measure of the place estimation error, as the token difference between a marking and its estimate in a given place.

Definition 11. *Let us consider a place $p \in P$ and an observed word $w \in L(N, M_0)$. Let M_w and μ_w be the corresponding marking and its estimate. The place estimation error in p is $e_p(M_w, \mu_w) = M_w(p) - \mu_w(p)$ and its update after the firing of t is $e_p(M_w, \mu'_{wt}) = M_w(p) - \mu'_{wt}(p)$.*

Analogously, it is possible to define a measure of the estimation error, as the token difference between a marking and its estimate.

Definition 12. *Given a marking M_w and its estimate μ_w , the estimation error is $e(M_w, \mu_w) = \sum_{p \in P} e_p(M_w, \mu_w) = \vec{1}_m^T \cdot (M_w - \mu_w)$ and its update after the firing of t is $e(M_w, \mu'_{wt}) = \vec{1}_m^T \cdot (M_w - \mu'_{wt})$.*

Note that the place estimation error is a monotonically non-increasing function of the observed word length.

Proposition 13. *Let $w = t_{\alpha_1} t_{\alpha_2} \dots \in L(N, M_0)$ be an observed word and w_i its prefix of length i . Then $\forall i$ and $\forall p$:*

$$e_p(M_{w_i}, \mu_{w_i}) \geq e_p(M_{w_i}, \mu'_{w_{i+1}}) = e_p(M_{w_{i+1}}, \mu_{w_{i+1}}), \quad (1)$$

and

$$e_p(M_{w_i}, \mu'_{w_{i+1}}) = \min \{e_p(M_{w_i}, \mu_{w_i}), M_{w_i} - \text{Pre}(p, t_{\alpha_{i+1}})\}. \quad (2)$$

Proof. To prove the first statement, we observe that by proposition 8, $\mu_{w_i}(p) \leq \mu'_{w_{i+1}}(p) \leq M_{w_i}(p)$, hence $e_p(M_{w_i}, \mu_{w_i}) \geq e_p(M_{w_i}, \mu'_{w_{i+1}})$. Also $e_p(M_{w_i}, \mu'_{w_{i+1}}) = (M_{w_i}(p) - \mu'_{w_{i+1}}(p)) = (M_{w_i}(p) + C(p, t_{\alpha_{i+1}}) - \mu'_{w_{i+1}}(p) - C(p, t_{\alpha_{i+1}})) = (M_{w_{i+1}}(p) - \mu_{w_{i+1}}(p)) = e_p(M_{w_{i+1}}, \mu_{w_{i+1}})$.

To prove the second statement, we observe that $e_p(M_{w_i}, \mu'_{w_{i+1}}) = M_{w_i}(p) - \mu'_{w_{i+1}}(p) = M_{w_i}(p) - \max\{\mu_{w_i}(p), \text{Pre}(p, t_{\alpha_{i+1}})\} = \min\{M_{w_i}(p) - \mu_{w_i}(p), M_{w_i}(p) - \text{Pre}(p, t_{\alpha_{i+1}})\} = \min\{e_p(M_{w_i}, \mu_{w_i}), M_{w_i}(p) - \text{Pre}(p, t_{\alpha_{i+1}})\}$. \square

Thus, it follows that also the estimation error is a monotonically non-increasing function of the observed word length.

Proposition 14. *Let $w = t_{\alpha_1}t_{\alpha_2}\dots \in L(N, M_0)$ be an observed word, w_i the prefix of w of length i , and μ_{w_i} and μ'_{w_i} the estimate and the updated estimate of M_{w_i} . Then $\forall i$:*

$$e(M_{w_i}, \mu_{w_i}) \geq e(M_{w_i}, \mu'_{w_{i+1}}) = e(M_{w_{i+1}}, \mu_{w_{i+1}}).$$

Proof. It immediately follows from proposition 13. \square

4 Properties of estimates

It is natural to ask under which conditions the estimated marking computed by algorithm 7 converges to the actual marking. This motivated us to define the following properties.

Definition 15. *Given a net system $\langle N, M_0 \rangle$, and a place $p \in P$, we say that a word $w \in L(N, M_0)$ is*

- *p -complete if $e_p(M_w, \mu_w) = 0$, i.e., if $\mu_w(p) = M_w(p)$;*
- *marking complete if w is p -complete for all $p \in P$.*

Thus a marking complete word allows one to reconstruct the actual marking of the net.

Based on this, we can define these properties of a net system.

Definition 16. *A net system $\langle N, M_0 \rangle$ is:*

- *marking observable (MO) if there exists a marking complete word $w \in L(N, M_0)$;*
- *strongly marking observable (SMO) in k steps if:*

1. $\forall w \in L(N, M_0)$ such that $|w| \geq k$, w is marking complete,
2. $\forall w \in L(N, M_0)$ such that $|w| < k$, either w is marking complete or $\exists t \in T$ such that $M_0[wt]$.

In the above definitions we note that the observability properties depend not only on the net structure N , but also on the initial marking M_0 , that we assume is unknown. Thus, it may seem that those properties have little significance per se. In effect, we will use the characterization of MO and SMO to define two more general properties that have greater significance.

Definition 17. *A net system $\langle N, M_0 \rangle$ is:*

- uniformly marking observable (uMO) *if $\forall M \in R(N, M_0)$, $\langle N, M \rangle$ is MO;*
- uniformly strongly marking observable (uSMO) *in k steps if $\forall M \in R(N, M_0)$, $\langle N, M \rangle$ is SMO in k steps.*

The property of uMO and uSMO are important if we consider the following problem: we consider a system whose initial marking M_0 is known. Due to a communication failure the system evolves unobserved. When the communication is re-established, we can only be sure that the actual marking belongs to the set $R(N, M_0)$. We want to know if the marking can be reconstructed starting from any of these reachable markings.

Definition 18. *A net N is:*

- structurally marking observable (sMO) *if it is MO for any initial marking $M_0 \in \mathbb{N}^m$;*
- structurally strongly marking observable (sSMO) *in k steps if $\langle N, M_0 \rangle$ is SMO in a number of steps k (that depends on $M_0 \in \mathbb{N}^m$).*

The properties of sMO and sSMO are even more general and only depend on the net structure N .

The above properties are related among them as shown in the following partial order diagram:

$$\begin{array}{ccccc}
 \text{sSMO} & \longrightarrow & \text{uSMO} & \longrightarrow & \text{SMO} \\
 \downarrow & & \downarrow & & \downarrow \\
 \text{sMO} & \longrightarrow & \text{uMO} & \longrightarrow & \text{MO}
 \end{array}$$

Here, say, $\text{sMO} \longrightarrow \text{uMO}$ means that if a net N is sMO then $\langle N, M_0 \rangle$ is uMO for all initial markings M_0 . By means of simple counterexamples it is possible to prove that properties not in a partial order relationship are uncorrelated.

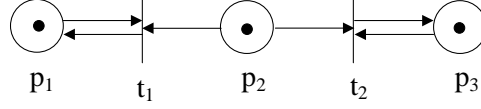


Figure 1: A net system that is not MO, but whose places are MO.

Note that sometimes, only the marking of a subset of places can be reconstructed, thus making it necessary to introduce the following definition.

Definition 19. Given a net system $\langle N, M_0 \rangle$, a place $p \in P$ is:

- marking observable (MO) if there exists a p -complete word $w \in L(N, M_0)$;
- strongly marking observable (SMO) in k_p steps (where k_p depends on the place p) if:
 1. $\forall w \in L(N, M_0)$ such that $|w| \geq k_p$, w is p -complete;
 2. $\forall w \in L(N, M_0)$ such that $|w| < k_p$, either w is p -complete or $\exists t \in T$ such that $M_0[wt]$.

Analogously, we can extend to a single place p all the properties of estimates previously defined for a net (system), namely, uMO, uSMO, sMO, sSMO. For sake of brevity we omit formal definitions. We highlight, however, the following implications.

- $\forall p$, p is MO $\iff \langle N, M_0 \rangle$ is MO
- $\forall p$, p is SMO (uMO, uSMO, sMO, sSMO) \iff

$$\langle N, M_0 \rangle \text{ is SMO (uMO, uSMO, sMO, sSMO).}$$

Note that the first one only holds in one sense. In fact, even if all places are observable, this does not imply that there exists one sequence that reconstructs the marking of all places.

Example 20. Let us consider the net system $\langle N, M_0 \rangle$ in figure 1. All places are MO but the net system is not MO. In fact, if t_1 fires, we reconstruct the marking of places p_1 and p_2 , but the net reaches a dead marking, thus making it impossible to reconstruct the marking of place p_3 . Analogously, the firing of t_2 enables us to reconstruct the actual marking of places p_2 and p_3 , but it produces a deadlock, thus not enabling the reconstruction of the marking in p_1 . ■

5 Observer coverability graph

In this subsection we show how to construct an *observer coverability tree* and the corresponding *observer coverability graph (OCG)* to represent both the set of reachable markings of a net system and the error of the estimate computed in accordance with algorithm 7. More precisely, each node of the OCG contains a vector M covering a marking of the net and an upper bound error vector $u \in \mathbb{N}^m$. We will show that the OCG is useful to prove observability properties.

Algorithm 21 (Observer coverability tree).

1. Let $u_0 = M_0$. Label the initial node (M_0/u_0) as the root and tag it "new".
2. If "new" nodes exist, select a new node (M/u) and:
 - 2.1. If (M/u) is identical to a node labeled "old" then tag (M/u) "old" and go to step 2.
 - 2.2. If no transitions are enabled at M , tag (M/u) "dead" and go to step 2.
 - 2.3. For each transition t enabled at M do the following:
 - 2.3.1. $\forall p \in P$, if $M(p) = \omega$ then let $\tilde{M}(p) = M(p)$ and $\tilde{u}(p) = u(p)$,
else let $\tilde{M}(p) = M(p) + C(p, t)$ and $\tilde{u}(p) = \min\{u(p), M(p) - Pre(p, t)\}$;
 - 2.3.2. on the path from the root to (M/u) if there exists a marking $\bar{M} \leq \tilde{M}$ and $\tilde{M} \neq \bar{M}$, i.e., \bar{M} is covered by \tilde{M} , then let $\tilde{M}(p) = \omega$ for each p such that $\tilde{M}(p) > \bar{M}(p)$;
 - 2.3.3. introduce (\tilde{M}/\tilde{u}) as a node, draw an arc with label t from (M/u) to (\tilde{M}/\tilde{u}) , and tag (\tilde{M}/\tilde{u}) "new".
 - 2.4 Tag (M/u) "old" and go to step 2. ■

Note that its construction follows the well known rules of a coverability tree for a P/T net [17]. Also, we note that the error bound vector u is set to the actual error for the root node and then it is updated as we add new nodes. Note, however, that whenever we reach a marking whose component $M(p)$ is ω , the error bound $u(p)$ is not updated any more (see step 2.3.1).

The **observer coverability graph** of a Petri net $\langle N, M_0 \rangle$ is a labeled directed graph $\mathcal{G} = (V, E)$ with transition function $\delta : V \times E \rightarrow V$. Its node set V is the set of all distinct labeled nodes in the observer coverability tree, and each arc in E is labeled with a transition t to represent a firing such that $\delta((M/u), t) = (M'/u')$, where (M/u) and (M'/u') are in V . Note that in the OCG all tags used in the construction of the observer coverability tree are omitted.

We will also represent the initial marking by a round corner box, while a thick box represents a marking whose estimation error bound vector is $u = \vec{0}_m$.

Example 22. Let us consider the net systems in figure 2 and their OCG. Since the two nets are unbounded, in both cases ω appears. The OCG of a bounded net is shown in figure 3. ■

Let us demonstrate that the OCG of a P/T net has a finite number of nodes.

Property 23. *Let \mathcal{G} be the OCG of $\langle N, M_0 \rangle$. The number of nodes in \mathcal{G} is bounded by $v = v' \cdot \prod_{p \in P} (M_0(p) + 1)$ where v' is the number of nodes in the usual coverability graph of $\langle N, M_0 \rangle$.*

Proof. By virtue of algorithm 21 the initial error bound vector is equal to the initial estimate, i.e., $u_0 = M_0$. Moreover, by proposition 14 the place estimation error is a monotonically non-increasing function of the observed word length, thus the estimation error in the generic place p may assume at most $M_0(p) + 1$ different values. It follows that the number of nodes in \mathcal{G} is limited by the number of nodes v' in the coverability graph times $\prod_{p \in P} (M_0(p) + 1)$. □

Apart from the well known properties that can be studied through the coverability graph of a P/T net [17], the OCG enables us to study some more interesting properties.

Proposition 24. *Let \mathcal{G} be the OCG of $\langle N, M_0 \rangle$. Given $w \in L(N, M_0)$ consider the node (M/u) reached on the graph executing w , i.e., let $(M/u) = \delta((M_0/u_0), w)$. It holds that:*

- (i) *the place estimation error $e_p(M_w, \mu_w) \in [\ell(p), u(p)]$ where $u(p)$ is the component of u corresponding to place p and*

$$\ell(p) = \begin{cases} u(p) & \text{if } M(p) \neq \omega \\ 0 & \text{if } M(p) = \omega \end{cases}$$

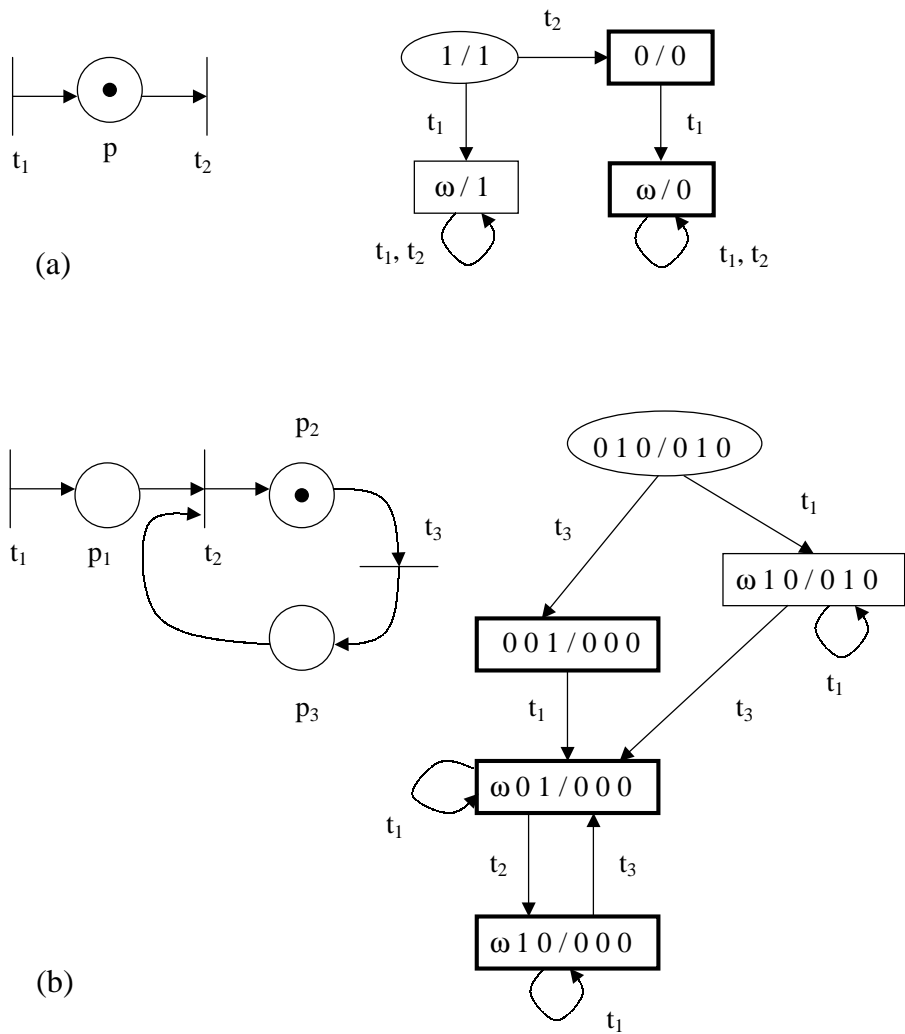


Figure 2: *Unbounded Petri nets and their observer coverability graphs.*

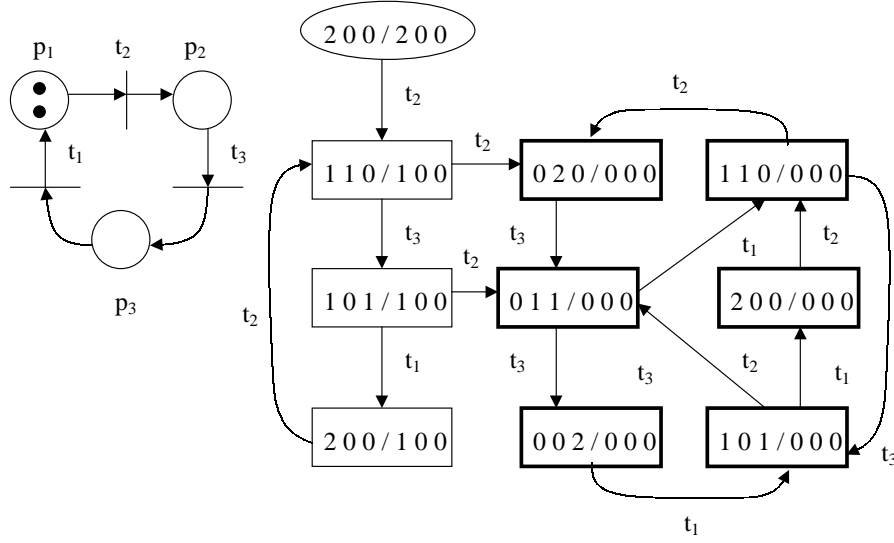


Figure 3: A bounded Petri net and its observer coverability graph.

(ii) the estimation error

$$e(M_w, \mu_w) \in \left[\sum_{p \in P} \ell(p), \sum_{p \in P} u(p) \right].$$

Proof. We prove this by induction on the length of w .

- (i) When $w \equiv w_0$, i.e., w is a word of null length, $(M/u) = (M_0/u_0)$, and $e_p(M_{w_0}, \mu_{w_0}) = M_0(p) - 0 = M_0(p) \equiv u_0(p)$.

Assume that the property (i) holds for a word $w' \in L(N, M_0)$ and let $\delta((M_0/u_0), w') = (M'/u')$. Let t be an enabled transition at $M_{w'}$ and $w = w't$: in the OCG there will be a transition $\delta((M'/u'), t) = (M/u)$. We can consider two cases.

If $M(p) \neq \omega$, then $M_{w'} = M'(p) \neq \omega$ and

$$\begin{aligned} e_p(M_w, \mu_w) &= \min\{e_p(M_{w'}, \mu_{w'}), M_{w'}(p) - Pre(p, t)\} \\ &= \min\{u'(p), M'(p) - Pre(p, t)\} = u(p). \end{aligned}$$

where the first equality derives from eq. (1), the second one from the induction hypothesis, and the third one from step 2.3.1 of algorithm 21.

If $M(p) = \omega$, then

$$\begin{aligned} e_p(M_w, \mu_w) &= \min\{e_p(M_{w'}, \mu_{w'}), M_{w'}(p) - Pre(p, t)\} \\ &\leq e_p(M_{w'}, \mu_{w'}) \leq u'(p) = u(p). \end{aligned}$$

where the last inequality derives from the induction hypothesis, and the last equality from step 2.3.1 of algorithm 21.

(ii) Immediately follows from the previous item. \square

Example 25. Let us consider again the net system in figure 2.a and its OCG. The estimation error relative to the node labeled with $(\omega/1)$ may be either null or unitary. If we consider $w = t_1 t_2 t_2$ then $e_p(M_w, \mu_w) = 0$, thus on the OCG we read an upper bound of the estimation error. On the contrary, $e_p(M_w, \mu_w) = 1$ is the exact estimation error for all words w such that $\forall w' \preceq w, |w'|_{t_1} \geq |w|_{t_2}$.

Now, let us consider the net system in figure 2.b. Here, every node with label $M(p_1) = \omega$ is also characterized by $u(p_1) = 0$, i.e., the upper bound on the place estimation error in p_1 is null. Therefore, in this case in each node of the OCG we can read the actual place estimation error in p_1 .

Finally, in the example in figure 3 no ω appears in \mathcal{G} being the net bounded, thus in each node u is the exact estimation error vector. \blacksquare

6 Properties analysis

In this subsection we discuss in detail the observability problem. In particular we provide necessary and sufficient conditions to characterize the properties defined above and we also prove that all these properties are decidable. The OCG is a useful tool when dealing with some analysis problems.

6.1 Word completeness

An elementary necessary and sufficient condition for completeness of a word is the following.

Proposition 26. A word $w \in L(N, M_0)$ is: (i) p -complete iff $M_0(p) = w^{-1}(\mu_w)(p)$; (ii) marking complete iff $M_0 = w^{-1}(\mu_w)$.

Proof. The word w is p -complete (resp., marking complete) if and only if $M_w(p) = \mu_w(p)$ (resp., $M_w = \mu_w$), i.e., if and only if $M_0(p) = w^{-1}(M_w)(p) = w^{-1}(\mu_w)(p)$ (resp., $M_0 = w^{-1}(M_w) = w^{-1}(\mu_w)$).

Example 27. Let us consider the net system in figure 3. The word $w = t_2$ is not marking complete because $\mu_{t_2} = [0 \ 1 \ 0]^T$ and $t_2^{-1}(\mu_{t_2}) = [1 \ 0 \ 0]^T \not\leq M_0$. On the contrary, the word $w = t_2 t_2$ is marking complete because $\mu_{t_2 t_2} = [0 \ 2 \ 0]^T$ and $[t_2 t_2]^{-1}(\mu_{t_2 t_2}) = [2 \ 0 \ 0]^T = M_0$. ■

Another semi-decision procedure for completeness can be given using the OCG.

Proposition 28. Let us consider a net system $\langle N, M_0 \rangle$ and its OCG \mathcal{G} . Let (M/u) be the node in \mathcal{G} reached executing $w \in L(N, M_0)$, i.e., $(M/u) = \delta((M_0/u_0), w)$ and let us consider a place $p \in P$.

(i) If $u(p) = 0$ (resp., $u = \vec{0}_m$), then w is p -complete (resp., marking complete).

(ii) If $M(p) \neq \omega$ and $u(p) \neq 0$, then w is not p -complete, hence it is not marking complete.

Proof: It follows from proposition 24. □

Note that the OCG provides necessary and sufficient conditions for the completeness of a word only in the case of bounded P/T nets, when ω does not appear in the graph. On the contrary, it only provides two distinct sufficient or necessary conditions for the completeness of a word in the case of unbounded nets.

Example 29. Let us consider again the bounded net system in figure 3. The OCG allows one to say that the word $w = t_2$ is not marking complete because its execution leads to $(1 \ 1 \ 0/1 \ 0 \ 0)$, while the word $w = t_2 t_2$ is marking complete because its execution leads to $(0 \ 2 \ 0/0 \ 0 \ 0)$.

Let us consider the unbounded net system in figure 2.a. If we consider $w = t_1 t_2 t_2$, w is complete but this is not deducible from the OCG because its execution leads to $(\omega/1)$. ■

Theorem 30. Let $\langle N, M_0 \rangle$ be a net system and w a word in $L(N, M_0)$. It is decidable whether w is marking complete wrt to $\langle N, M_0 \rangle$.

Proof. It follows from proposition 26, because it is sufficient to determine μ_w using algorithm 7,

and then compute $w^{-1}(\mu_w)$. □

6.2 Observability

We firstly provide a necessary and sufficient condition for marking observability.

Proposition 31. *The net system $\langle N, M_0 \rangle$ is marking observable iff*

$$L(N, M_0) \supseteq \bigcup_{\overline{M}_0 \preceq M_0} L(N, \overline{M}_0).$$

Proof. In general $L(N, M_0) \supseteq \bigcup_{\overline{M}_0 \preceq M_0} L(N, \overline{M}_0)$. We prove that the system is not observable iff the equality holds.

In fact $L(N, M_0) = \bigcup_{\overline{M}_0 \preceq M_0} L(N, \overline{M}_0) \iff \forall w \in L(N, M_0), \exists \overline{M}_0 \preceq M_0$ such that $w \in L(N, \overline{M}_0) \iff$ (by lemma 9) $\forall w \in L(N, M_0), \exists \overline{M}_0$ such that $M_0 \succeq \overline{M}_0 \geq w^{-1}(\mu_w) \iff$ (by proposition 26) $\forall w \in L(N, M_0), w$ is not complete $\iff \langle N, M_0 \rangle$ is not marking observable. □

Checking for language inclusion is difficult (see [20]) thus we look for simpler decision procedures. In particular the OCG provides a simpler semi-decision (i.e., only sufficient) condition for the marking observability.

Proposition 32. *Let us consider a net system $\langle N, M_0 \rangle$ and its OCG \mathcal{G} . (i) A place p is marking observable if there exists a node in \mathcal{G} such that $u(p) = 0$. (ii) The net system is marking observable if there exists a node in \mathcal{G} such that $u = \vec{0}_m$.*

Proof: It follows from the definition of marking observability and from proposition 28. □

On the contrary, the OCG provides necessary and sufficient conditions for strong marking observability. Let us first demonstrate, as an intermediate result, that the repeated firing of a repetitive sequence does not decrease the estimation error.

Lemma 33. *Let $\langle N, M_0 \rangle$ be a net system and let us assume that there exists a firing sequence w' that enables a repetitive sequence w , i.e., $M_0[w']M_{w'}[w]M_{w'w}$ with $M_{w'w} \geq M_{w'}$. Then $\forall p \in P$ and $\forall i > 1$, $e_p(M_{w'^iw^i}, \mu_{w'^iw^i}) = e_p(M_{w'w}, \mu_{w'w})$.*

Proof. While observing a sequence w , the error may decrease only during step 4 of algorithm 7, i.e., when we compute the updating estimate.

Let t be the first transition in the sequence w . If t fires after $w'w^i$, in step 4 of algorithm 7 we have

$$\mu'_{w'w^i t} \geq \text{Pre}(\cdot, t).$$

Using proposition 13 it is easy to show that for all $i \geq 1$

$$(M_{w'w^{i+1}} - \mu_{w'w^{i+1}}) \leq (M_{w'w^i} - \mu'_{w'w^i t}).$$

thus

$$\mu_{w'w^{i+1}} \geq (M_{w'w^{i+1}} - M_{w'w^i}) + \mu'_{w'w^i t} \geq \mu'_{w'w^i t} \geq \text{Pre}(\cdot, t).$$

Therefore, $\mu'_{w'w^{i+1}t} = \mu_{w'w^{i+1}}$, i.e., the estimate is not updated and the error remains constant each time w is repeated after it has fired once. \square

Proposition 34. *Let us consider a net system $\langle N, M_0 \rangle$, its OCG \mathcal{G} , and a place p of N . The place p (resp., the net system) is strongly marking observable in k_p steps iff the error bound vector is such that $u(p) = 0$ (resp., $u = \vec{0}_m$) for each node (M/u) in \mathcal{G} such that: (a) the node (M/u) is in a cycle; (b) the node (M/u) is dead. Moreover, if (a) and (b) are satisfied, it is possible to compute k_p as the length¹ of the longest directed path that starts from the root, contains only intermediate nodes with $u(p) > 0$ (resp., $u \succeq \vec{0}_m$), and ends on a node with $u(p) = 0$ (resp., $u = \vec{0}_m$).*

Proof: We prove the property for a single place p ; the proof for the net system trivially follows.

(if) By proposition 23, the number of nodes in \mathcal{G} is finite and equal to v . Thus any word w of length greater or equal to v must pass through a cycle in \mathcal{G} , hence w is p -complete by assumption (a). Any word of length less than v that leads to a dead marking is also p -complete, by assumption (b). This is sufficient to show that the place is SMO in k_p steps with $k_p \leq v$. The actual value of k_p may be computed as suggested in the statement.

(only if) We show this by contradiction, proving that if any of the two conditions are violated the place cannot be SMO. Clearly, if condition (b) is violated, the place is not strongly marking observable by definition. Now, let assume that (a) is violated. We consider two subcases.

¹The length of a path is given by the number of edges along the path.

(i) Assume there exists a node (M/u) along a cycle γ of \mathcal{G} with $M(p) \neq \omega$ and $u(p) > 0$. Then there exists w' such that $M_{w'} = M$ and $e_p(M_{w'}, \mu_{w'}) = u(p) > 0$. The cycle γ corresponds to a word w such that $M_{w'}[w]M_{w'}$, i.e., by Lemma 33 the infinite length sequence $w'w^i$ may be fired for all $i > 0$ without reducing the estimation error and the place is not SMO.

(ii) Assume there exists a node (M/u) with $M(p) = \omega$ and $u(p) > 0$ (we do not even need to assume it is along a cycle). Then consider the path along the observer coverability tree that reaches (M/u) from (M_0/u_0) and let (\tilde{M}, \tilde{u}) be the first node we encounter along this path with $M(p) = \omega$. Then, at step 2.3.2 of algorithm 21, we have identified a marking \overline{M} such that $M_0[w']\overline{M}[w]M_{w'w}$ and $M_{w'w} \geq M_{w'}$ (\tilde{M} is obtained from $M_{w'w}$ by changing in ω the components greater than the corresponding components of \overline{M}). Also, $e_p(M_{w'w}, \mu_{w'w}) = \tilde{u}(p) \geq u(p) > 0$. Thus, by Lemma 33 the infinite length sequence $w'w^i$ may be fired for all $i > 0$ without reducing the estimation error and the place is not SMO. \square

Example 35. All net systems in figures 2–3 are MO but not SMO. On the contrary, one example of strong marking observability (in one step) can be obtained if we consider the net in figure 3 with initial marking $M_0 = [1 \ 0 \ 0]^T$ (see figure 4.a).

Analogously, the net systems in figure 4.c–d are MO, but only the net system in figure 4.c is SMO (in one step). In the case of the net system in figure 4.d, there exist arbitrarily long sequences that are enabled at the initial marking, and that are not complete. In fact, $\forall i \in \mathbb{N}$, $w \in t_1(t_2)^i$ is not complete.

These results can also be read in table 1 that summarizes all observability properties of P/T nets in figure 4. \blacksquare

Finally, let us discuss the decidability of these properties.

Theorem 36. *It is decidable whether the net system $\langle N, M_0 \rangle$ is marking and strongly marking observable.*

Proof. Decidability of the MO property follows from propositions 31 and the fact that language inclusion for free Petri nets languages is decidable [8, 20]. Decidability of the SMO property follows from proposition 34, since the OCG is finite. \square

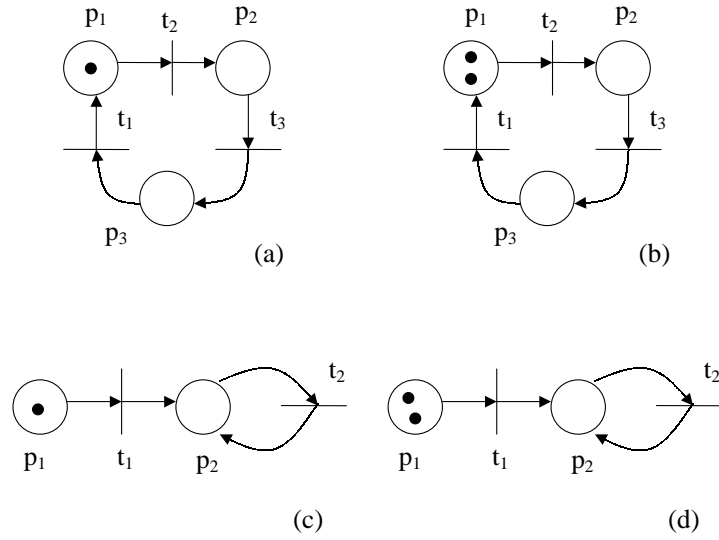


Figure 4: *Place/Transition nets used in the examples.*

Table 1: *Observability properties of the nets in figure 4.*

6.3 Uniform observability

In this section we first provide necessary and sufficient conditions for both uniform MO and uniform SMO. Then we prove the decidability of both these properties.

Let us first demonstrate an important lemma.

Lemma 37. *Let $\langle N, M_0 \rangle$ be a net system. A place $p \in P$ is observable in $\langle N, M_0 \rangle$ iff at least one element in the semi-linear set*

$$\mathcal{A}_p = \{M \in \mathbb{N}^m \mid M(p) = 0\} \cup \left(\bigcup_{t \in p^\bullet} \{M \in \mathbb{N}^m \mid M(p) = \text{Pre}(p, t), M \geq \text{Pre}(\cdot, t)\} \right) \quad (3)$$

is reachable.

Proof. (if) Let w be a word in $L(N, M_0)$. Let us consider two subcases.

- i) If $M_w \in \{M \in \mathbb{N}^m \mid M(p) = 0\}$, then $0 = M_w(p) \geq \mu_w(p) \geq 0$, thus $M_w(p) = \mu_w(p)$.
- ii) If $M_w \in \{M \in \mathbb{N}^m \mid M(p) = \text{Pre}(p, t), M \geq \text{Pre}(\cdot, t)\}$ where $t \in p^\bullet$, then t may fire at M_w and since $M_w(p) = \text{Pre}(p, t)$ the updated estimate is $\mu'_w(p) = M_w(p)$, hence $M_{wt}(p) = \mu_{wt}(p)$.

(only if) We prove this by contradiction.

If no marking with $M(p) = 0$ is reachable, then $M_w(p) > 0 \forall w \in L(N, M_0)$, thus the initial place estimation error is strictly positive. It may decrease only during step 4 of algorithm 7. However, if $\forall w$ and $\forall t \in p^\bullet$, $M_w(p) > \text{Pre}(p, t)$, then $\mu'_w(p) < M_w(p)$, thus the place estimation error keeps positive. \square

By virtue of the previous lemma, the study of uniform marking observability reduces to the study of m home space problems.

Proposition 38. *A net system $\langle N, M_0 \rangle$ is uniformly marking observable iff the semi-linear set \mathcal{A}_p given by eq. (3) is a home space $\forall p \in P$.*

Proof. It follows from the previous lemma and the fact that a net system $\langle N, M_0 \rangle$ is uniformly marking observable iff each place $p \in P$ is observable in $\langle N, M \rangle$, $\forall M \in R(N, M_0)$, i.e., iff the semi-linear set (3) is a home-space $\forall p \in P$. \square

Let us now consider the uniform SMO property.

Proposition 39. *A net system $\langle N, M_0 \rangle$ is uniformly strongly marking observable only if its reachability set is finite.*

Proof. If the reachability set is not finite, then (by fact 1.iii) there exist words w' and w such that $M_0[w']M_{w'}[w]M_{w'w}$ with $M_{w'w} \succeq M_{w'}$. This means that $w \in L(N, M_{w'}) \cap L(N, M_{w'w})$, and (by lemma 9) $M_{w'w} \not\prec M_{w'} \geq w^{-1}(\mu_w)$, thus (by proposition 26) the word w is not marking complete wrt $\langle N, M_{w'w} \rangle$. Also, we can have words of infinite length w^i (for all $i > 1$) that are not marking complete (by lemma 33) thus the system $\langle N, M_{w'w} \rangle$ is not SMO and finally $\langle N, M_0 \rangle$ is not uSMO. \square

Example 40. The net systems in figure 4.a, b and c are uMO. In fact, in the first two cases, $\forall p \in P$ the set $\{M \in \mathbb{N}^m \mid M(p) = 0\}$ is always a home space. In the third case, the set $\{M \in \mathbb{N}^m \mid M(p_1) = 0\}$ and the set $\{M \in \mathbb{N}^m \mid M(p_2) = Pre(p_2, t_2) = 1, M \geq Pre(\cdot, t_2)\}$ are home spaces. On the contrary, the net system in figure 4.d is not uMO because $\{M \in \mathbb{N}^m \mid M(p_2) = 0\} \cup \{M \in \mathbb{N}^m \mid M(p_2) = Pre(p_2, t_2) = 1, M \geq Pre(\cdot, t_2)\}$ is not a home space and thus p_2 is not uMO; in fact the net system $\langle N, M \rangle$ is not MO at $M = [0 \ 2]^T \in R(N, M_0)$.

The net systems in figure 4.a and c are uSMO. Obviously, the net system in figure 4.b is not uSMO, being not SMO. Analogously, being the net system in figure 4.d not uMO, it is also not uSMO. \blacksquare

Theorem 41. *It is decidable if a net system $\langle N, M_0 \rangle$ is uniformly and strongly uniformly marking observable.*

Proof. Let us first prove the decidability of uniform marking observability. Because of proposition 38 it is sufficient to prove that the home-space property for the set \mathcal{A}_p is decidable. Let us observe that $\forall p \in P$ the semi-linear set in eq. (3) is given by the finite union of linear sets having the same periods. In fact, if we consider a generic place $p_k \in P$,

$$\{M \in \mathbb{N}^m \mid M(p_k) = 0\} = \left\{ \sum_{i \neq k} a_i \vec{\varepsilon}_i \mid a_i \in \mathbb{N} \right\}$$

$$\{M \in \mathbb{N}^m \mid M(p_k) = Pre(p_k, t), M \geq Pre(\cdot, t)\} =$$

$$\left\{ Pre(\cdot, t) + \sum_{i \neq k} b_i \vec{\varepsilon}_i \mid b_i \in \mathbb{N} \right\}$$

where $\vec{\varepsilon}_i$ is the i -th canonical basis vector of dimension m . Thus, the decidability of the home-space property for \mathcal{A}_p immediately follows by theorem 5.

Secondly, let us prove the decidability of strong uniform marking observability. Let us observe that if the necessary requirement stated by proposition 39 is satisfied, then the reachability set is finite and the uniform strong marking observability can be verified by proving the strong marking observability — that is decidable [9] — for a finite set of initial markings. \square

6.4 Structural observability

In this subsection we provide necessary and sufficient conditions for both structural and strong structural marking observability and we prove the decidability of these properties.

Proving structural observability, requires the study of the system properties for all possible initial markings. Next two lemmas show that to prove that a place is observable for all initial markings in \mathbb{N}^m , just a finite subset of \mathbb{N}^m needs to be checked.

Lemma 42. *If a place $p \in P$ is observable in $\langle N, M \rangle$ then it is also observable in $\langle N, \overline{M} \rangle$ $\forall \overline{M} \geq M$ with $\overline{M}(p) = M(p)$.*

Proof. A place p is observable in $\langle N, M \rangle$ if and only if $\exists w \in L(N, M)$ such that $M[w]M'$ and $\mu_w(p) = M'(p)$. In this case $\forall \overline{M} \geq M$ with $\overline{M}(p) = M(p)$, $w \in L(N, \overline{M})$ (by fact 1.i) and $\overline{M}[w]\overline{M}'$ with $\overline{M}'(p) = M'(p) = \mu_w(p)$ (by fact 1.ii), i.e., p is also observable in $\langle N, \overline{M} \rangle$. \square

Lemma 43. *Let N be a Petri net and let $r_p = \max_{t \in T} \text{Pre}(p, t)$. Let*

$$M_i^p = \begin{cases} M_i^p(p') = 0 & \text{if } p' \neq p \\ M_i^p(p) = i. \end{cases} \quad (4)$$

A place $p \in P$ is observable in $\langle N, M_i^p \rangle \forall i \in \mathbb{N}$, iff p is observable in $\langle N, M_i^p \rangle$ for $i = 1, \dots, r_p + 1$.

Proof. If p is observable in $\langle N, M_{r_p+1}^p \rangle$ then $\exists w$ and $t \in p^\bullet$ such that $M_{r_p+1}^p[w]\overline{M}$, $\overline{M}(p) = \text{Pre}(p, t)$, i.e., the firing of the word w reduces the number of tokens in p . This implies that for all M_i^p with $i > r_p + 1$ the word w may also fire until we reach a marking \overline{M}' such that $\overline{M}' \geq M_{\varrho_i}^p$ and $\overline{M}'(p) = \varrho_i \leq r_p$. Since p is observable in $\langle N, M_{\varrho_i}^p \rangle$, then it is also observable in $\langle N, \overline{M}' \rangle$ by lemma 42. \square

Proposition 44. *A Petri net N is structurally marking observable iff $\forall p \in P$, p is observable in $\langle N, M_i^p \rangle$, where M_i^p is defined as in equation (4) and $i = 1, \dots, r_p + 1$.*

Proof. By definition a Petri net N is sMO iff $\forall p \in P$, p is observable in $\langle N, M \rangle \forall M \in \mathbb{N}^m$. By lemma 42 and lemma 43 it is however sufficient to check that each p is observable for the finite number of initial makings given in the statement. \square

Proposition 45. *A Petri net N is strongly structurally marking observable iff*

(a) *N has no repetitive sequences;*

(b) *$\forall p \in P, \exists t \in T$ such that*

$$Pre(p', t) = \begin{cases} 1 & \text{if } p' = p \\ 0 & \text{if } p' \neq p \end{cases}.$$

Proof. (if) We will prove that (a) and (b) imply that for any initial marking M_0 in finite number of steps the net loses all its tokens: this is a sufficient condition for SMO of $\langle N, M_0 \rangle$ by lemma 37. In fact, if no repetitive sequences exist, for any initial markings the length of all words firable is bounded, i.e., after a finite number of firings the net reaches a dead marking. Furthermore, if assumption (b) is verified, for each place $p \in P$ there exists a transition t whose single input is p and the corresponding arc weight is unitary, i.e., if t cannot fire then place p must be empty. Thus the dead marking must be the zero marking.

(only if) We prove this by contradiction.

Let us first assume that (a) is violated, and let w be a repetitive sequence. Clearly, for any $M \succeq w^{-1}(\mu_w)$, the word w is not marking complete wrt $\langle N, M \rangle$ (by proposition 26). Also, we can have words of infinite length w^i (for all $i > 1$) that are not marking complete (by lemma 33) thus the system $\langle N, M \rangle$ is not SMO.

Secondly, we assume that (a) is verified while $\exists p \in P$ such (b) is violated. We first observe that we can exclude the existence of transitions with no input arcs, because this would violate condition (a). Then it is obvious that given the marking M_1^p as in equation (4) (that contains one token in p and zero tokens elsewhere) no transition is enabled, thus the marking of p cannot be observed. \square

Theorem 46. *It is decidable if a Petri net N is structurally marking observable and structurally strongly marking observable.*

Proof. To prove that N is sMO it is sufficient to prove (by propositions 44) that all places are observable in $\langle N, M_0 \rangle$ for a finite number of initial markings M_0 . The property of being observable for a place is decidable because of theorem 3 and of the characterization given by lemma 37.

To prove that N is sSMO it is sufficient to check by propositions 45 that no repetitive sequences exist (and this may be checked with linear algebraic tools given the net incidence matrix) and that the net structure satisfies condition (b) (this may be checked by inspection). □

Example 47. Being sMO and sSMO structural properties of the net, the same conclusions can be drawn for nets in figure 2.a, b and c, d, respectively.

In particular, the net in figure 4.a is sMO by proposition 44. On the contrary, it is not sSMO. In fact, if we consider the initial marking in figure 4.b the net system is not SMO.

The net N in figure 4.c is not sMO (thus it is also not sSMO). In fact, if we consider $M_0 = [0 \ 2]^T$, $\langle N, M_0 \rangle$ is not MO. ■

A final remark regards the classes of nets that are sSMO. Although this property is rather easy to prove, the class of nets that satisfy this property is of little practical interest (they must become empty and deadlock in a finite number of steps). The property of structural MO, on the contrary, is more difficult to prove but is satisfied by a wider (more interesting) class of nets.

7 Marking estimation with macromarking

In the previous sections no information was assumed on the initial marking M_0 that originates the observed transition firings. It is often the case, however, that partial information about this marking is available.

As an example, let us assume that the net starts its evolution at a given time instant τ_- from a known marking M_- (called start marking). After having evolved unobserved for

some time, the net reaches a marking M_0 (called initial marking) from which we begin the observation of the transition firings. Now we know that $M_0 \in R(N, M_-)$ and we could use this information to better characterize the set of markings consistent with an observed word w given the information on the start marking as:

$$\mathcal{M}(w \mid M_-) = \{M \mid \exists M_0 \in R(N, M_-), M_0[w]M\}.$$

The main problem with this is that this characterization is given in terms of Petri net reachability (the initial marking must be reachable from the start marking) that is hard to solve. Looking for simpler structures, we consider the case in which the knowledge of M_0 can be written as $M_0 \in \mathcal{V}(V, \vec{b})$, where \mathcal{V} is a macromarking defined as follows.

Definition 48. Assume the set of places P can be written as the union of $r + 1$ subsets $P = P_0 \cup P_1 \cup \dots \cup P_r$, where $P_0 \cap P_j = \emptyset$ for all $j > 0$, while any two sets P_j and $P_{j'}$ may have a non null intersection if $j, j' > 0$. For each P_j , \vec{v}_j is its characteristic vector (i.e., $v_j(p) = 1$ if $p \in P_j$, else $v_j(p) = 0$). The number of tokens in P_0 is unknown. Let $V = [\vec{v}_1, \dots, \vec{v}_r]$ and $\vec{b} = [b_1, \dots, b_r]$. The macromarking $\mathcal{V}(V, \vec{b})$ is defined as the set $\{M \in \mathbb{N}^m \mid V^T M = \vec{b}\}$.

Note that, as a special case, if V is a matrix of P-invariants, then by definition

$$R(N, M_-) \subseteq \{M \in \mathbb{N}^m \mid V^T M = V^T M_-\}$$

where $V^T M_- = \vec{b}$ is known, thus a macromarking can also approximate the info about the start marking M_- .

We thus add to assumptions A1 and A2 given in section 3, the following assumption:

A3) the initial marking M_0 belongs to the macromarking $\mathcal{V}(V, \vec{b})$, i.e., it satisfies the equation

$$V^T M_0 = \vec{b}.$$

Given an evolution of the net $M_0[t_{\alpha_1}]M_1[t_{\alpha_2}] \dots$, we use the following algorithm to compute estimate μ_{w_i} and bound B_{w_i} of each actual marking M_{w_i} based on the observation of the word of events $w_i = t_{\alpha_1} t_{\alpha_2} \dots t_{\alpha_i}$, and of the knowledge of the initial macromarking $\mathcal{V}(V, \vec{b})$.

Algorithm 49. *Marking Estimation with Event Observation and Initial Macromarking*

1. Let the initial estimate be $\mu_{w_0} = \vec{0}_m$.
2. Let the initial bound be $B_{w_0} = \vec{b}$.

3. Let $i = 1$.
4. Wait until t_{α_i} fires.
5. Update the estimate $\mu_{w_{i-1}}$ to μ'_{w_i} with

$$\mu'_{w_i}(p) = \max\{\mu_{w_{i-1}}(p), \text{Pre}(p, t_{\alpha_i})\}.$$

6. Let $\mu_{w_i} = \mu'_{w_i} + C(\cdot, t_{\alpha_i})$.
7. Let $B_{w_i} = B_{w_{i-1}} - V^T \cdot (\mu'_{w_i} - \mu_{w_{i-1}})$.
8. Let $i = i + 1$.
9. Goto 4. ■

Note that the estimate μ computed using this algorithm is the same of the estimate computed with Algorithm 7 and thus, all observability properties already discussed do not change.

What is new is the additional information given by the bounds that will be used to characterize the set of consistent markings.

Theorem 50. *Given an observed word $w \in L(N, M_0)$ with initial macromarking $\mathcal{V}(V, \vec{b})$, the corresponding estimated marking μ_w and bound B_w computed by Algorithm 49, the set of consistent markings is $\mathcal{M}(w \mid V, \vec{b}) = \{M \in \mathbb{N}^n \mid V^T \cdot M = V^T \cdot \mu_w + B_w, M \geq \mu_w\}$.*

Proof. Let w_0 be the empty word. Then $\{M \in \mathbb{N}^n \mid V^T \cdot M = V^T \cdot \mu_{w_0} + B_{w_0}, M \geq \mu_{w_0}\} = \{M \in \mathbb{N}^n \mid V^T \cdot M = V^T \cdot \mu_{w_0} + \vec{b} - V^T \cdot \mu_{w_0}, M \geq \mu_{w_0}\} = \{M \in \mathbb{N}^n \mid V^T \cdot M = \vec{b}, M \geq \mu_{w_0}\} = \{M \in \mathbb{N}^n \mid V^T \cdot M = \vec{b}\} \equiv \mathcal{M}(w_0 \mid V, \vec{b})$.

By induction, let us show that $\mathcal{M}(w \mid V, \vec{b}) = \{M \in \mathbb{N}^n \mid M \geq \mu_w, V^T \cdot M = V^T \cdot \mu_w + B_w\} \implies \mathcal{M}(wt \mid V, \vec{b}) = \{M \in \mathbb{N}^n \mid M \geq \mu_{wt}, V^T \cdot M = V^T \cdot \mu_{wt} + B_{wt}\}$.

In fact $\mathcal{M}(wt \mid V, \vec{b}) \equiv \{M \in \mathbb{N}^n \mid \exists M' \in \mathcal{M}(w \mid V, \vec{b}), M' \geq \text{Pre}(\cdot, t), M = M' + C(\cdot, t)\} = \{M \in \mathbb{N}^n \mid \exists M', V^T \cdot M' = V^T \cdot \mu_w + B_w, M' \geq \mu_w, M' \geq \text{Pre}(\cdot, t), M = M' + C(\cdot, t)\}$.

Now, let μ'_{wt} be the updated estimate of M_w after t fires. Then $[M' \geq \mu_w] \wedge [M' \geq \text{Pre}(\cdot, t)] \iff M' \geq \mu'_{wt}$. Furthermore, with the notation of Algorithm 49, $B_w + V^T \cdot \mu_w = B_{wt} + V^T \cdot \mu'_{wt}$, and $\mu_{wt} = \mu'_{wt} + C(\cdot, t)$. Hence $\mathcal{M}(wt \mid V, \vec{b}) = \{M \in \mathbb{N}^n \mid \exists M', V^T \cdot M' = V^T \cdot \mu'_{wt} + B_{wt}, M' \geq \mu'_{wt}, M = M' + C(\cdot, t)\} = \{M \in \mathbb{N}^n \mid V^T \cdot M = V^T \cdot \mu_{wt} + B_{wt}, M \geq \mu_{wt}\}$. \square

The previous theorem allows us to write a general optimization problem of the form

$$\begin{aligned} & \max \vec{c}^T \cdot M \\ \text{s.t. } & M \in \mathcal{M}(w \mid V, \vec{b}) \end{aligned}$$

as a linear integer programming problem (IPP)

$$\begin{aligned} & \max \vec{c}^T \cdot M \\ \text{s.t. } & V^T \cdot M = V^T \cdot \mu_w + B_w \\ & M \geq \mu_w. \end{aligned} \tag{5}$$

As an example, appropriately choosing the value of \vec{c} , such an IPP can be used to compute the maximum over all consistent markings of the tokens in the net (if $\vec{c} = \vec{1}$), and of the tokens in a generic place p_i (if $\vec{c} = \vec{e}_i$).

Note that if we do not want to solve an integer linear programming problem, it is possible to give ranges on the estimation errors by simple inspection of B .

Theorem 51. *Consider an observed word $w \in L(N, M_0)$ with initial macromarking $\mathcal{V}(V, \vec{b})$, the corresponding estimated marking μ_w and bound B_w computed by Algorithm 49, and $P = P_0 \cup P_1 \cup \dots \cup P_r$ with the notation of Definition 48.*

1. $\forall M \in \mathcal{M}(w \mid V, \vec{b})$, $l \leq e(M, \mu_w) \leq u$ where $l = \max_j B_w(j)$, and $u = \vec{1}_r^T \cdot B_w$ if $P_0 = \emptyset$, else $u = +\infty$.
2. $\forall M \in \mathcal{M}(w \mid V, \vec{b})$, $e_{p_i}(M, \mu_w) \leq u_{p_i}$ where $u_{p_i} = \min_{j \mid p_i \in P_j} B_w(j)$ if $p_i \in P \setminus P_0$, else $u_{p_i} = +\infty$.

Proof. 1. The first inequality immediately follows by definition of B_w . Moreover, $V^T M = V^T \mu_w + B_w$, thus $V^T(M - \mu_w) = B_w$ and $\vec{1}_r^T \cdot V^T(M - \mu_w) = \vec{1}_r^T \cdot B_w$. If $P_0 = \emptyset$, then $\vec{1}_r^T \cdot V^T \geq \vec{1}_m^T$, thus $e(M, \mu_w) = \vec{1}_m^T \cdot (M - \mu_w) \leq \vec{1}_r^T \cdot V^T(M - \mu_w) = \vec{1}_r^T \cdot B_w$. On the contrary, if $P_0 \neq \emptyset$, then an arbitrarily large number of tokens can be added to P_0 , thus $u = +\infty$.

2. By definition, each subset P_j such that $p_i \in P_j$ imposes a constraint of the form $e_{p_i}(M, \mu_w) \leq B_w(j)$ on the place estimation error. When p_i belongs to more than

one subset of places, the resulting constraints should be satisfied simultaneously, thus providing the above statement. On the contrary, if $p_i \in P_0$, no limit exists on the number of tokens that can be added to μ_w .

□

Remark 52. *In the case of disjoint subsets P_j 's, $l = u = \vec{1}_r^T \cdot B_w$ if $P_0 = \emptyset$, else $l = \vec{1}_r^T \cdot B_w$.*

From Theorem 51, we have the following corollary that shows how the bound B_w may be used to prove that a word w is complete.

Corollary 53. *With the notation of Algorithm 49:*

1. *if $P_0 = \emptyset$, then w is marking complete if and only if $B_w = \vec{0}_r$;*
2. *if $P_0 \neq \emptyset$, then w is marking complete only if $B_w = \vec{0}_r$.*

We conclude this section with the following observation. In Algorithm 49 by construction we are sure that for all w it holds $\mu_w \leq M_w$. However, if we are willing to pay the extra cost of solving m optimization problems of the form (5) at each iteration, we may be able to update each component of μ_w in step 1 and step 6 of the algorithm as follows:

$$\tilde{\mu}_w(p) = \min\{M(p) \mid M \in \mathcal{M}(w \mid V, \vec{b})\}.$$

It is easy to show that this updated estimate is such that

$$\mu_w \leq \tilde{\mu}_w \leq M_w.$$

8 Control using observers

In this section we show how the marking estimate constructed with the formalism discussed in the previous section can be used by a control agent to enforce a given specification on the plant behaviour.

We make several assumptions that are briefly discussed here.

- The specification is given as a set of forbidden markings \mathcal{F} . The set of legal markings is $\mathcal{L} = \mathbb{N}^m - \mathcal{F}$.

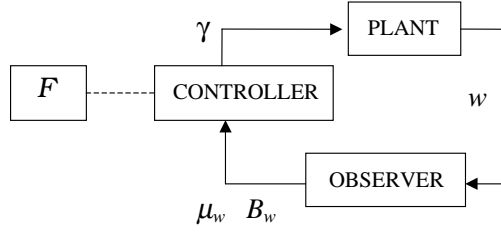


Figure 5: *State feedback control loop with observer.*

- The controller may disable transitions to prevent the plant from entering a forbidden marking. From the knowledge of μ_w and B_w , the controller computes a control pattern $\gamma : T \rightarrow \{0, 1\}$. If $\gamma(t) = 0$ then t is disabled by the controller.
- All transitions are controllable, i.e., can be disabled by the controller.

The considered control scheme is shown in Figure 5.

Under the assumption that the initial marking $M_0 \in \mathcal{L}$, the following algorithm may be used by the controller at each step to ensure that markings in \mathcal{F} are not reached.

Algorithm 54. *Let w be the observed word, and $\mathcal{M}(w \mid V, \vec{b}) = \{M \in \mathbb{N}^n \mid V^T \cdot M = V^T \cdot \mu_w + B_w, M \geq \mu_w\}$, where μ_w and B_w are computed by the observer.*

for all $t \in T$

begin

$\gamma(t) := 1;$

if $\exists M \in \mathcal{M}(w \mid V, \vec{b}) \cap \mathcal{L}$ such that $M[t]M', M' \in \mathcal{F}$

then $\gamma(t) := 0;$

end.

■

Clearly this algorithm prevents all transition firings that lead from \mathcal{L} to \mathcal{F} but is not necessarily optimal, in the sense that it may also prevent transition firings that lead from \mathcal{L} to \mathcal{L} . A similar algorithm was also discussed in [11] (Algorithm 5.3) to ensure predicate invariance using state estimates computed by a dynamic observer.

In general, it may be difficult to check the condition of the **if** statement of the algorithm. However, when \mathcal{F} is a finite set the observer estimate may be used to verify this condition.

In fact, we simply have to check whether there exists a marking $M \in \mathcal{M}(wt \mid V, \vec{b}) \cap \mathcal{F}$ such that $t^{-1}(M) \in \mathcal{L}$, and this is trivial given the characterization of Theorem 50.

We also would like to consider a special case in which a control law different from the one presented above may be suitable. Let the specification on the legal states be given by $\mathcal{L} = \{M \in \mathbb{N}^m \mid W^T \cdot M \leq \vec{k}\}$ where $W = [\vec{w}_1 \cdots \vec{w}_q]$ with $\vec{w}_j \in \mathbb{Z}^n$ and $\vec{k} = [k_1 \cdots k_q]$ with $k_j \in \mathbb{Z}$. This kind of specifications, that we call *generalized mutual exclusion constraints* have been considered by various authors [10, 12, 24].

Assume that the initial marking M_0 of the plant does not necessarily belong to \mathcal{L} (this is a natural assumption when considering error recovery problems). Then, given a marking M we may want to prevent the firing of transition t such that $M[t]M'$ when both these two conditions are verified:

- (a) there exists \vec{w}_j with $\vec{w}_j \cdot M' > k_j$, i.e., $M' \in \mathcal{F}$;
- (b) $\vec{w}_j \cdot M' > \vec{w}_j \cdot M$, i.e., the firing of t either leads to a violation of the constraint (if $M \in \mathcal{L}$) or to a "worse" violation of the constraint (if $M \in \mathcal{F}$).

In this case the following algorithm may be used to compute the control pattern γ at each step.

Algorithm 55. *Let w be the observed word, and $\mathcal{M}(w \mid V, \vec{b}) = \{M \in \mathbb{N}^n \mid V^T \cdot M = V^T \cdot \mu_w + B_w, M \geq \mu_w\}$, where μ_w and B_w are computed by the observer. Let $\mathcal{L} = \{M \in \mathbb{N}^n \mid W^T \cdot M \leq \vec{k}\}$.*

for all $t \in T$

begin

$\gamma(t) := 1$;

$j := 1$;

while $j \leq q$ **and** $\gamma(t) = 1$ **do**

begin

$\Delta := \vec{w}_j^T \cdot C(\cdot, t)$;

if $\Delta > 0$ **then**

begin

```

 $\bar{m} := \max \{ \vec{w}_j^T \cdot M \mid M \in \mathcal{M}(wt \mid V, \vec{b}) \};$ 
if  $\bar{m} > k_j$  then  $\gamma(t) := 0;$ 
end;
 $j := j + 1;$ 
end;
end.

```

■

Thus a transition is disabled at M only if its firing leads to a marking M' such that for at least one constraint j : $\vec{w}_j^T \cdot M' > \vec{w}_j^T \cdot M$ (i.e., $\Delta > 0$) and there exists a consistent marking M'' in $\mathcal{M}(wt \mid V, \vec{b})$ that violates the constraint (i.e., $\vec{w}_j^T \cdot M'' > k_j$).

The methodology developed in this paper is applied to a simple manufacturing example. Other examples can be found in [7].

Example 56. Let us consider the net in Figure 3 with initial marking $M_0 = [1 \ 1 \ 1]^T$. This system may represent a pool of three machines. Each token represents a machine that may be in any of three states: working (token in place p_1), idle (token in place p_2), loading (token in place p_3). We assume that the specification on the system behavior requires that at most two machines may be simultaneously working, i.e., the set of forbidden states is $\mathcal{F} = \{M \in \mathbb{N}^3 \mid M(p_1) > 2\}$.

The initial macromarking $M(p_1) + M(p_2) + M(p_3) = 3$ captures our knowledge that there are three machines in the pool. Their initial state is, however, unknown.

To represent the global behavior of the plant with observer under control using Algorithm 55, we have represented the observer reachability graph of the controlled plant with observer in Figure 6. The observer reachability graph has been constructed following the same rules of Algorithm 21. We have also introduced a new label at each node so as to better highlight the effect of the control pattern γ . Each node is now labeled $(M/u/B)$ where M is the real marking, u is a vector whose components, being the net bounded, coincide with the place estimation errors, and B is the resulting bound.

Note that, given $u = M - \mu$, the bound B can be immediately computed as $B = V^T u$,

thus the addition of the new label does not introduce significant variations on Algorithm 21.

Let us briefly discuss the graph in Figure 6. The initial marking is represented by a round corner box. A dashed box represents a marking that cannot be reached because the transition firing leading to it is disabled by the controller (the corresponding edge is dashed). A thick box represents a marking reached by a complete word w , i.e., $u_w = \vec{0}$ and $B_w = 0$: the future evolution from such a marking is not shown.

Note that the transition firings disabled by the controller using Algorithm 55 in reality do not lead to forbidden markings: they are disabled because there exist markings consistent with the observation from which these transition firings would lead to forbidden markings. This can be easily verified by looking at the nodes within dashed boxes. In all these cases the value of \bar{m} in Algorithm 55 is equal to

$$\bar{m} = V^T \cdot \mu + B = V^T \cdot (M - u) + B = M(p_1) - u(p_1) + B = 3.$$

On the contrary, if the real marking would have been used to determine the control pattern, such a node would have been reachable, being $V^T \cdot M \leq 2$. ■

Let us finally observe that, since the controller may prevent the firing of transitions whose firing is perfectly legal, it may also be the case that the controlled system is blocking.

A preliminary solution to this problem has been presented in [7] and consists in the introduction of suitable recovery mechanisms with an "ad hoc" reasoning. A more general procedure to automatically recover the net from a blocking condition is given in [1]. This approach is essentially based on a linear algebraic characterization of deadlock markings, that reveal to be useful to derive additional information on the actual marking of the net, so as to improve the marking estimate, thus restricting the set of w consistent markings.

9 Conclusions

In this paper we dealt with the problem of estimating the marking of a Place/Transition net based on event observation, assuming that the net structure is known. We considered two cases: (a) the initial marking is not known; (b) the initial marking is known to belong to a *macromarking*, i.e., we know the token contents of subsets of places but not the exact token

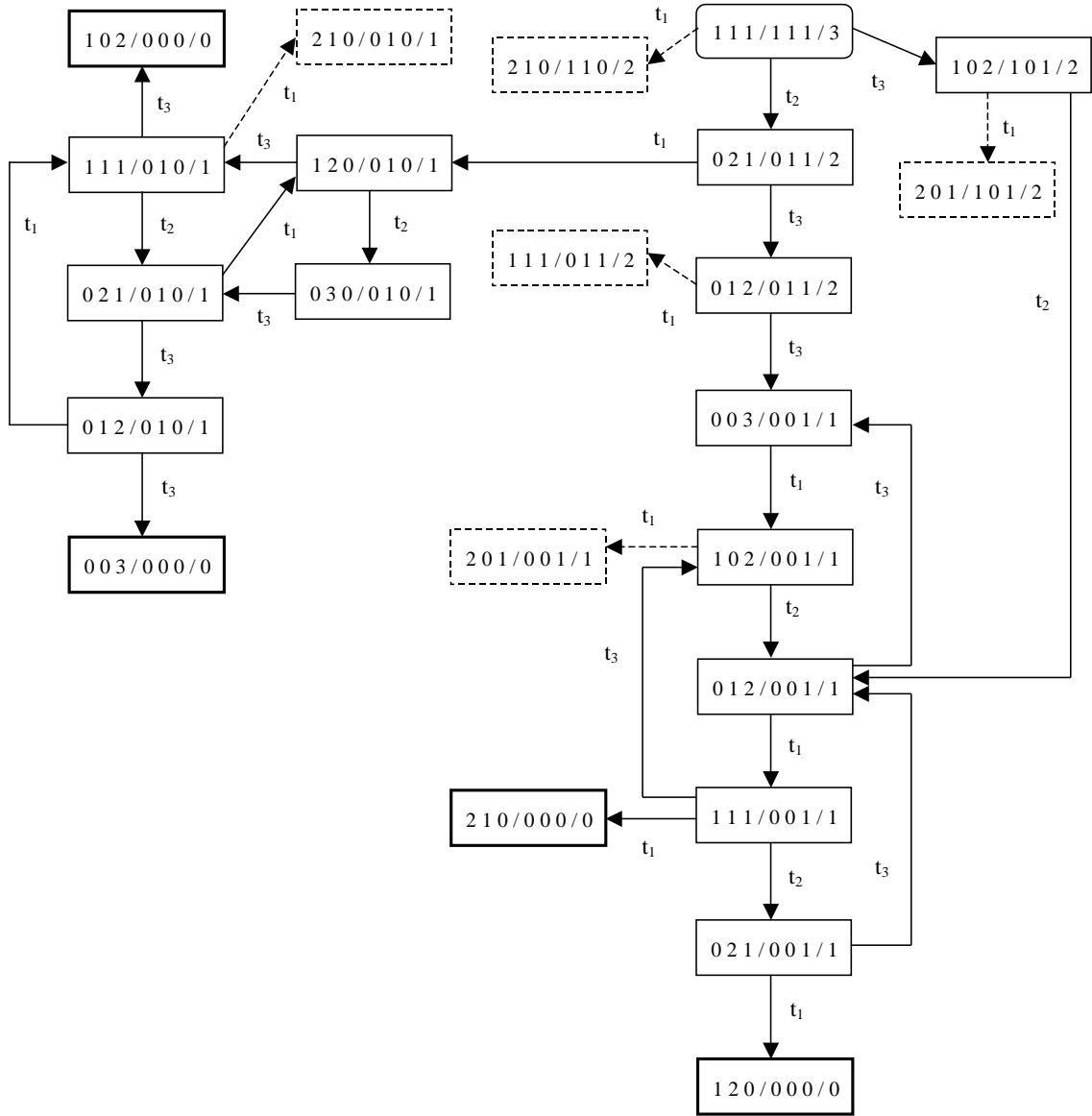


Figure 6: *Observer reachability graph of the controlled net system in example 56.*

location.

We defined several observability properties and showed that they are decidable. In particular we considered two main properties: *marking observability* and *strong marking observability*. The first one means that there exists at least one word that is complete, while the second one means that all words can be completed in a finite number of steps to a complete word.

We investigated the possibility that the two properties above are satisfied by a net N starting from an initial marking M_0 , by a net N starting from any marking M reachable from an initial marking M_0 (*uniform observability*) or by a net N starting from any marking in \mathbb{N}^m (*structural observability*) where m is the number of places of the net.

We also introduced the *observer coverability graph*, i.e., the usual coverability graph of a P/T net augmented with a vector that keeps track of the estimation error on each place of the net. We proved that it can be a useful tool when proving some of the above properties. We also showed that many observability properties can be proved by reducing them to other decision problems (e.g., home-space properties, marking reachability, existence of repetitive sequences) that can be checked using algorithms well known from the literature.

Finally, we showed how the estimate generated by the observer may be used to design a state feedback controller, that ensures that the controlled system never enters a set of forbidden states.

There are several ways in which this research may be extended.

Firstly, we may consider the case in which not all transition firings are observable, or there may be transitions that do not generate distinct events. This may destroy the framework used in this paper because in this case the marking estimate is not anymore a lower bound of the actual marking. However, in some restricted case, e.g., when the firing of all unobservable transitions does not change the total number of tokens, we believe it may be possible to extend the approach used in this paper.

Secondly, we may associate a probabilistic structure to the transition firings. Then, given an initial marking M_0 , we may define a function $\Pi : \mathbb{N} \rightarrow [0, 1]$: $\Pi(k)$ denotes the probability that, after having observed k event firings, we obtain a complete word. It would be interesting to study under which conditions the limit of $\Pi(k)$ goes to 1 as k goes to infinity.

References

- [1] F. Basile, P. Chiacchio, A. Giua, C. Seatzu, “Deadlock recovery of controlled Petri net models using observers,” *8th IEEE Int. Conf. on Emerging Technologies and Factory Automation* (Antibes, France), pp. 441–449, Oct. 2001.
- [2] M.S. Bazaraa, J.J. Jarvis, *Linear Programming and Network Flows*, John Wiley & sons, 1977.
- [3] P.E. Caines, R. Greiner, S. Wang, “Dynamical Logic Observers for Finite Automata,” *Proc. 27th Conf. on Decision and Control* (Austin, Texas), pp. 226–233, Dec. 1988.
- [4] P.E. Caines, S. Wang, “Classical and Logic Based Regulator Design and its Complexity for Partially Observed Automata,” *Proc. 28th Int. Conf. on Decision and Control* (Tampa, Florida), pp. 132–137, Dec. 1989.
- [5] J. Cardoso, R. Valette, D. Dubois, “Petri Nets with Uncertain Markings,” *Advances in Petri Nets 1990*, G. Rozenberg (ed.), LNCS Vol. 483, pp. 64–78, Springer-Verlag, 1991.
- [6] C. Johnen, D. Frutos Escrig, “Decidability of home space property,” *Report LRI 503*, Univ. d’Orsay, RI, June 1989.
- [7] A. Fanni, A. Giua, N. Sanna, “Control and Error Recovery of Petri Net Models with Event Observers,” *Proc. 2nd Int. Work. on Manufacturing and Petri Nets* (Toulouse, France), pp. 53–68, June 1997.
- [8] A. Giua, F. DiCesare, “Decidability and Closure Properties of Weak Petri Net Languages in Supervisory Control,” *IEEE Trans. on Automatic Control*, Vol. 40, No. 5, pp. 906–910, 1995.
- [9] A. Giua, “Petri Net State Estimators Based on Event Observation,” *Proc. 36th Int. Conf. on Decision and Control* (San Diego, California), pp. 4086–4091, December 1997.
- [10] A. Giua, F. DiCesare, M. Silva, “Generalized Mutual Exclusion Constraints on Nets with Uncontrollable Transitions,” *Proc. 1992 IEEE Int. Conf. on Systems, Man, and Cybernetics* (Chicago, Illinois), pp. 974–979, Oct. 1992.

- [11] R. Kumar, V. Garg, S.I. Markus, “Predicates and Predicate Transformers for Supervisory Control of Discrete Event Dynamical Systems,” *IEEE Trans. on Automatic Control*, Vol. 38, No. 2, pp. 232–247, 1993.
- [12] Y. Li, W.M. Wonham, “Control of Vector Discrete-Event Systems — Part II: Controller Synthesis,” *IEEE Trans. on Automatic Control*, Vol. 39, No. 3, pp. 512–531, 1994.
- [13] Y. Li, W.M. Wonham, “Controllability and Observability in the State-Feedback Control of Discrete-Event Systems,” *Proc. 27th Conf. on Decision and Control* (Austin, Texas), pp. 203–207, Dec. 1988.
- [14] Y. Li, W.M. Wonham, “Control of Vector Discrete-Event Systems — Part I: The Base Model,” *IEEE Trans. on Automatic Control*, Vol. 38, No. 8, pp. 1215–1227, 1993.
- [15] M.E. Meda, A. Ramírez, A. Malo, “Identification in Discrete Event Systems,” *Proc. IEEE Int. Conf. on Systems, Man and Cybernetics*, San Diego, CA, pp. 740–5, Oct. 1998.
- [16] G. Memmi, J. Vautherin, “Analysing Nets by the Invariant Method,” In: *Petri Nets: Central Models and their Properties*, W. Brauer, W. Reisig, G. Rozenberg, LNCS 254, Springer Verlag, pp. 300–336, 1987.
- [17] T. Murata, “Petri Nets: Properties, Analysis and Applications,” *Proceedings IEEE*, Vol. 77, No. 4, pp. 541–580, 1989.
- [18] C.M. Özveren, A.S. Willsky, “Observability of discrete event dynamic systems,” *IEEE Trans. on Automatic Control*, Vol. 35, No. 7, pp. 797–806, July 1990.
- [19] R.J. Parikh, “Language Generating Devices,” *MIT, Research laboratory of Electronics, Quarterly Progress Report 60*, pp. 191–212, 1961.
- [20] E. Pelz, “Closure Properties of Deterministic Petri Net Languages,” *Proc. STACS 1987*, LNCS No. 247, pp. 373–382, Springer-Verlag, 1987.
- [21] P.J. Ramadge, “Observability of Discrete-Event Systems,” *Proc. 25th Conf. on Decision and Control* (Athens, Greece), pp. 1108–1112, Dec. 1986.

- [22] A. Ramírez-Treviño, I. Rivera-Rangel, E. López-Mellado, “Observer Design for Discrete Event Systems modeled by Interpreted Petri Nets,” *2000 IEEE Int. Conf. on Robotics and Automation*, pp. 2871–2876, April 2000.
- [23] S. Takai, T. Ushio, S. Kodama, “Static-State Feedback Control of Discrete-Event Systems Under Partial Observation,” *IEEE Trans. on Automatic Control*, Vol. 40, No. 11, pp. 1950–1955, 1995.
- [24] K. Yamalidou, J.O. Moody, M.D. Lemmon, P.J. Antsaklis, “Feedback Control of Petri Nets Based on Place Invariants,” *Automatica*, Vol. 32, No. 1, 1996.
- [25] C. Wang, M. Schwartz, “Fault Detection with Multiple Observers,” *IEEE/ACM Trans. on Networking*, Vol. 1, No. 1, pp. 48–55, 1993.
- [26] L. Zhang, L.E. Holloway, “Forbidden State Avoidance in Controlled Petri Nets Under Partial Observation,” *Proc. 33rd Allerton Conf.* (Monticello, Illinois), pp. 146–155, Oct. 1995.

List of changes

We are grateful to the editor and to the reviewers for their constructive critics and their encouragement. We have thoroughly revised the paper taking all their remarks into account.

Reply to the Editor

As we discuss in detail in the replies to each reviewer, we have addressed all three main points mentioned by the editor.

- We used the notion of minimal initial marking enabling a sequence w to simplify the characterization of a complete word given in Proposition 26, and the proof of Theorem 30 (was Theorem 29 in the original version) as suggested by reviewer 1.
- We have added a formal definition of home space in subsection 2.2, as suggested by reviewer 2.
- We have corrected the definition of "structurally strongly marking observable" as suggested by reviewer 3. Note, however, that all results in section 6 are not modified by this new definition.

There was also another series of minor changes that we made to make the paper more general and consistent. In the original version, the proof of some net properties were based on the corresponding place properties, while other were not. In effect, in a purely logical framework such as the one considered in the paper, the marking estimation in a subset of places may be more important than the estimation of the total marking. To this end, we have extended in the revised manuscript the notion of completeness with respect to a single place (see section 4). The presentation of the material in section 4 and 6 has also been partially modified so that the properties of completeness of an observed word and of observability of a net system are now derived from the more basic property of place observability.

1. In section 4 we have made these changes.
 - (a) The definition of a complete word (Definition 15) now also contains the notion of p -complete.

- (b) We have given more emphasis to the the definition of place observability (adding also strong observability for a place) in Definition 19.
- (c) We have added Example 20 and Figure 1.

2. In section 6 we have made these changes.

- (a) The characterization of a complete word using the initial marking in Proposition 26 and using the OCG in Proposition 28 now also contains the notion of p -completeness.
- (b) The characterization of observability using the OCG in Proposition 32 now also contains the notion of p -observability.
- (c) Lemma 33 is now given for a single place.
- (d) The characterization of strong observability using the OCG in Proposition 34 now also contains the notion of strong p -observability.
- (e) We have added Figure 4 and Table 1 and examples (see Example 35 and Example 40).

Finally, we have better rephrased Fact 1 (grouping fact 1 and 2 into a single one, and adding a third one), Proposition 39, and Definition 48.

Reply to Reviewer 1

Following the suggestion of reviewer 1 we have rephrased Lemma 9 and added a short comment just before it, showing that the minimal initial marking enabling a sequence w on a net N is $w^{-1}(\mu_w)$. Thanks to this, the characterization of a complete word given in Proposition 26, and the proof of Theorem 30 (was Theorem 29 in the original version) is much simpler.

All typos pointed out by the reviewer have been corrected.

Reply to Reviewer 2

The reviewer pointed out that the concept of *home space* has been used without explicitly defining it. We have added a formal definition of home space in subsection 2.2.

Reply to Reviewer 3

This is how we have addressed all the issues raised by the reviewer.

Issues raised in the “Comments to Authors” section

- The reviewer claims that “The first part of the paper, which deals with the case when the initial marking is known, is not very relevant for designing an observer, since all transitions are observable. The structural results do not seem to rely much on the results which assume the initial marking known.”

In section 4 we have clearly stated that the definitions of *marking observability* and *strong marking observability* have little significance per se. In effect, we use these characterization to “define two more general observability properties”, i.e., uniform and structural observability. (see comment before Definition 17).

- The reviewer claims that “Considering that all possible markings have the same probability may often be unrealistic. The probability information is always important for estimation”.

We agree with the reviewer’s consideration. In fact, even if the approach presented in the manuscript is purely logical, we have suggested as a direction for future research the idea of associating probabilities to the transition firings. In this way, the likelihood of having reached a given marking can also be computed.

- The reviewer says that “... it would be useful to comment on the existence of other than algorithm 5 observers which would not be subject to negative results.”

The comment of the reviewer originates from the fact that in section 8 we have shown that: (a) the marking estimate constructed with the formalism proposed in the paper can be used by a control agent to enforce a given specification on the plant behaviour; (b) the use of the estimate, instead of the real marking, may lead to the disabling of some transition firings that in reality do not yield forbidden markings, and this may lead the closed loop system to a blocking condition.

This problem does not originate from the particular estimation algorithm proposed in the paper, but is intrinsic in the problem itself (incomplete information).

We have mentioned at the end of section 8 that we are currently investigating this problem and a procedure that may automatically recover the net from a blocking condition is given in

- F. Basile, P. Chiacchio, A. Giua, C. Seatzu, “Deadlock recovery of controlled Petri net models using observers,” *8th IEEE Int. Conf. on Emerging Technologies and Factory Automation*, (Antibes, France), Oct. 2001 (submitted).

Issues raised in the “Technical Correctness” section

- The term between parenthesis in equation (3) is essential for the validity of the lemma (in the revised version it becomes Lemma 37). To clarify this point, we have discussed in Example 40 the case of the net system in figure 4.c, that is shown to be uMO because the set $\{M \in \mathbb{N}^m \mid M(p_1) = 0\}$ and the set $\{M \in \mathbb{N}^m \mid M(p_2) = Pre(p_2, t_2) = 1, M \geq Pre(\cdot, t_2)\}$ are home spaces.
- F1 and F2 are two typos: by F1 and F2 we meant Fact 1 and Fact 2 (in the revised version we grouped them in fact 1 in section 2.1). We have corrected the lemma (Lemma 42 of the revised version).

Issues raised in the “Quality of presentation” section

- We have added a formal definition of home space in subsection 2.2.

Issues raised in the “Other comments” section

- The definition that we gave for “structurally strongly marking observable” was not well formulated, and the correct definition we had in mind was that suggested by the reviewer. Thus we have restated the definition of sSMO (see Definition 18) as follows: “a net system is *structurally strongly marking observable* in k steps if $\langle N, M_0 \rangle$ is SMO in a number of steps k that depends on $M_0 \in \mathbb{N}^m$ ”.

Note, however, that all results in section 6 are not modified by this new definition.

- We thanks the reviewer for his/her useful suggestion for future research, that we have mentioned in the last section: "Firstly, we may consider the case in which not all transition firings are observable, or there may be transitions that do not generate distinct events".