Deadlock recovery of Petri net models controlled using observers

Francesco Basile, Pasquale Chiacchio
Dip. di Informatica e Sistemistica, Università degli Studi di Napoli Federico II, Italy
Phone: +39-081-768-3178 – Fax: +39-081-768-3186 – Email: {fbasile,chiacchio}@unina.it

Alessandro Giua[1], Carla Seatzu
Dip. di Ing. Elettrica ed Elettronica, Università di Cagliari, Italy
Phone: +39-070-675-5892 – Fax: +39-070-675-5900 – Email: {giua,seatzu}@diee.unica.it

Abstract — **This paper discusses the problem of controlling a Petri net whose marking cannot be measured but is estimated using an observer. The control objective is that of enforcing a set of generalized mutual exclusion constraints (GMEC) and all transitions are assumed to be controllable. 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 and it may also be the case that, as a result of this, the controlled system reaches a deadlock. We present a general approach, based on siphon analysis, to recover from such an "observer induced" deadlock. The most interesting feature of our approach is that the observer, controller and deadlock recovery algorithms are all based on the same linear algebraic techniques, thus allowing the overall problem to be solved using a single formalism.**

I. INTRODUCTION

This paper discusses the problem of controlling a Petri net whose marking cannot be measured.

When the structure and the initial marking of a 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 the cases in which the initial marking is know to belong to a "macromarking", i.e., we know the token contents of subsets of places but not the exact token distribution.

In [7] it was shown 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 an algorithm was given for computing the marking estimate μ_w and error bound B_w. 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 $\mathcal{M}$ of markings consistent with an observed word, i.e., the set of marking in which the system may actually be given the observed word, can easily be described in terms of the observer estimate and can be characterized as the integer solutions of a linear constraint set.

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. We consider a special class of specifications that limit the weighted sum of markings in subset of places called generalized mutual exclusion constraints (GMEC) [6]

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 enters a forbidden marking, the controller may prevent the firing of transitions whose firing is perfectly legal given the actual marking of the plant.

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 [17] 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 [9, 10] and by Takai *et al.* [15]. 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 mentioned 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. In particular, we consider the case in which, due to the incomplete knowledge of the plant state, the controlled system reaches a *blocking state*, i.e., a *deadlock* , and present a general approach, based on siphon analysis, to recover from such a blocking.

One the two main results of the paper consists in showing that the set of deadlock markings $\mathcal{M}_d$ of a structurally bounded net can be characterized as the integer solutions of a linear constraint set (see also the related work of [2]).

If we assume that no transition firing occurs within a reasonable amount of time in a controlled system, a deadlock can be detected. The second main result of this paper consist in

[1]Corresponding author

0-7803-7241-7/01/$10.00 (c)2001 IEEE

showing how the observer can use this information to restrict the set of consistent markings to $\mathcal{M}' = \mathcal{M} \cap \mathcal{M}_d$ and possibly recover from the blocking.

The paper is structured as follows. In Section II the notation on Petri nets is briefly recalled. In Section III the marking estimation problem is defined and an algorithm to design a state observer taken from [7] is given. In Section IV the control problem using observers is presented, and an example taken from the manufacturing domain is given to show that the presence of an observer in the feedback loop may easily lead to a deadlock. In Section V we present a linear algebraic characterization of deadlocks marking. In Section VI we show how the linear program derived in the previous section can be used by a deadlock recovery procedure; an example of application is also given.

II. BACKGROUND ON PETRI NETS

In this section we recall the formalism used in the paper. For more details on Petri nets we address to [12].

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; $Pre: P \times T \to \mathbb{N}$ and $Post: P \times T \to \mathbb{N}$ are the *pre–* and *post–* incidence functions that specify the arcs; $C = Post - Pre$ is the incidence matrix. The *preset* and *postset* of a place p are respectively: ${}^\bullet p = \{t \in T \mid Post(p,t) > 0\}$ and $p^\bullet = \{t \in T \mid Pre(p,t) > 0\}$.

A *marking* is a vector $M: P \to \mathbb{N}$ that assigns to each place of a P/T net a non–negative integer number of tokens, represented by black dots. In the following we denote as $M(p)$ 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 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 firable 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 $L(N, M_0)$.

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 nonnegative integer vector $x \neq \vec{0}_m$ such that $x^T C = \vec{0}_n{}^T$ is called a *P–invariant* (here $\vec{0}_k$ denotes a $k \times 1$ vector of zeros). A P-invariant is *minimal* if there does not exist a P-invariant y such that $y \leq x$.

A transition t is said to be *live* if for any $M \in R(N, M_0)$, there exists a sequence of transitions firable from M which contains t. A Petri net is said to be live if all transitions are *live*. A Petri net is said to be *deadlock–free* if at least one transition is enabled at every reachable marking.

A place p is said to be *bounded* if there exists a constant k such that $M(p) \leq k$ for all $M \in R(N, M_0)$. A net system is bounded if all places are bounded. A net is *structurally bounded* if it is bounded for all initial markings.

A P/T net is called *ordinary* when all of its arc weights are 1's. A *siphon* of an ordinary net is a *non–empty* set of places $\mathcal{S} \subseteq P$ such that: $\bigcup_{p \in \mathcal{S}} {}^\bullet p \subseteq \bigcup_{p \in \mathcal{S}} p^\bullet$. A siphon is *minimal* if it is not the superset of any other siphon. In the following we denote as $\vec{s} \in \{0,1\}^m$ the characteristic vector of $\mathcal{S}$, where $s_i = 1$ if place $p_i \in \mathcal{S}$ and $s_i = 0$ otherwise.

III. MARKING ESTIMATION WITH MACROMARKINGS

In previous works [7, 8] the authors dealt with the problem of reconstructing the marking of a P/T net, under the assumption that the structure of the net is known and the transition firings can be observed, but no information is available on the initial marking. An estimation algorithm has been proposed and several interesting observability properties have been investigated.

In this paper we assume that partial information about the initial marking is available. In particular, we assume that the initial marking is given in the form of a *macromarking*.

Definition 1 ([8]) *Assume the set of places P can be written as the union of $r+1$ subsets $P = P_0 \cup P_1 \cup \cdots \cup P_r$, 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., $\vec{v}_j(p) = 1$ if $p \in P_j$, else $\vec{v}_j(p) = 0$). The number of tokens contained in P_j ($j > 0$) is known to be b_j, while the number of tokens in P_0 is unknown.*

Let $V = [\vec{v}_1, \cdots, \vec{v}_r]$ and $\vec{b} = [b_1, b_2, \cdots, b_r]$. The macromarking *$\mathcal{V}(V, \vec{b})$ is defined as the set $\{M \in \mathbb{N}^n \mid V^T M = \vec{b}\}$.*

We make 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.

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

We also introduce the following notation.

Definition 2 ([7]) *After the word w has been observed we define the set $\mathcal{M}(w \mid V, \vec{b})$ of $w-$consistent markings as the set of all markings in which the system may be given the observed behaviour and the initial marking, i.e., the set*

$$\mathcal{M}(w \mid V, \vec{b}) = \{M \in \mathbb{N}^m \mid \exists M_0 \in \mathbb{N}^m,\ M_0[w\rangle M\}.$$

The use of macromarkings comes out quite naturally. 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 the set of markings consistent with an observed word w given the information on the start marking 1s:

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

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 $M_0 \in \mathcal{V}(V, \vec{b})$.

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_-.

The notion of macromarking also occurs frequently when describing systems containing a known set of resources (e.g., parts, machines) whose actual conditions (e.g., exact location of parts within the plant, state of a machine) is unknown [1].

Given an evolution of the net $M_0[t_{\alpha_1}\rangle M_1[t_{\alpha_2}\rangle \cdots$, 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 $w_i = t_{\alpha_1}, t_{\alpha_2}, \cdots, t_{\alpha_i}$, and of the knowledge of the initial macromarking $\mathcal{V}(V, \vec{b})$.

Algorithm 3 ([7]) *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

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

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. ■

In [8], it has been proved that the estimate computed using algorithm 3 is a lower bound on the actual marking of the net.

Proposition 4 ([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*

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

In [8] we have also defined a meaningful measure of the place estimation error, as the token difference between a marking and its estimate in a given place.

Definition 5 ([8]) *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 $e_p(M_w, \mu'_{wt}) = M_w(p) - \mu'_{wt}(p)$.*

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

Proposition 6 ([8]) *Let $w = t_{\alpha_1} t_{\alpha_2} \cdots \in L(N, M_0)$ be an observed word and w_i its prefix of length i. Then $\forall i$ and $\forall p$:*

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

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

The set of consistent markings can be characterized as follows.

Theorem 7 ([7]) *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 3, the set of* $w-$consistent markings *is*

$$\mathcal{M}(w \mid V, \vec{b}) = \{ \begin{array}{l} M \in \mathbb{N}^n \mid M \geq \mu_w, \\ V^T \cdot M = V^T \cdot \mu_w + B_w\}. \end{array}$$

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

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

as a linear integer programming problem (IPP)

$$\begin{cases} \max \vec{c}^T \cdot M \\ \text{s.t.} \\ V^T \cdot M = V^T \cdot \mu_w + B_w \\ M \geq \mu_w. \end{cases} \tag{2}$$

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 estimation error (if $\vec{c} = \vec{1}$), and of the error in a generic place p_i (if $\vec{c} = \vec{\varepsilon}_i$, where $\vec{\varepsilon}_i$ denotes the i–th canonical basis vector).

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 8 ([8]) *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 3, and $P = P_0 \cup P_1 \cup \cdots \cup P_r$ with the notation of Definition 1.*

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 $u_{p_i} = +\infty$.*

Remark 9 ([8]) *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 8, we have the following corollary that shows how the bound B_w may be used to prove that a word w is complete.

Corollary 10 ([8]) *With the notation of Algorithm 3:*

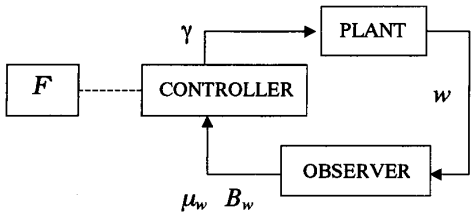

Fig. 1 State feedback control loop with observer.

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 3 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 an optimization problem of the form (2) 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

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

IV. 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 [5, 8].

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 $\mathcal{L} = \mathbb{N}^m - \mathcal{F}$.
- 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 $\gamma : T \to \{0, 1\}$. 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 Fig. 1.

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 11 *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* $M[t\rangle M'$, $M' \in \mathcal{F}$
    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 [4] (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 7.

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 $\vec{w}_j \in \mathbb{Z}^m$ 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 [6, 11, 16].

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 $M[t\rangle M'$ 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 12 *Let* w *be the observed word, and* $\mathcal{M}(w \mid V, \vec{b}) = \{M \in \mathbb{N}^m \mid V^T \cdot M = V^T \cdot \mu_w + B_w, M \geq \mu_w\}$, *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
      $\overline{m} := \max\{\vec{w}_j^T \cdot M \mid M \in \mathcal{M}(wt \mid V, \vec{b})\}$;
      if $\overline{m} > k_j$ **then** $\gamma(t) := 0$;
      end;
    $j := j + 1$;
    end;
  end.

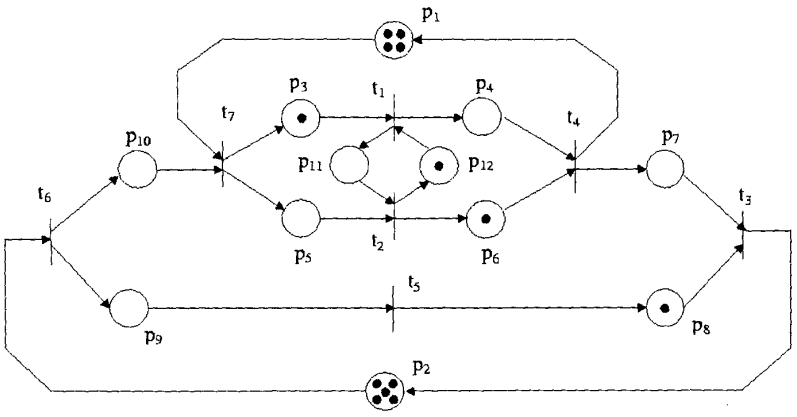

Fig. 2 Event graph model of the assembly system.

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., $\vec{w}_j^T \cdot M'' > k_j$).

A. A manufacturing example

Now, let us apply the above methodology to a manufacturing system whose Petri net model is shown in Fig 2.

This assembly system, that is similar to the one described in [3], consists of five machines, $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_3$, $\mathcal{M}_4$ and $\mathcal{M}_5$ whose operational process is modeled by the firing of transitions t_1, t_2, t_3, t_4 and t_5, respectively. Two principal types of operations are involved in this manufacturing system: *regular operations* and *assembly operations*. Regular operations (modeled by transitions t_1, t_2 and t_5) just transform a component of the intermediate product. Assembly operations (modeled by transitions t_3 and t_4) put components together to obtain a more complex component of a final product or the final product itself. Note that this model uses transitions (t_6 and t_7) which do not represent operations but the beginning of the manufacturing of components which are required to assemble a more complex component or the final product. In this example there are two manufacturing levels, the primary one, performed by $\mathcal{M}_3$, leads to finite product, the secondary one, performed by $\mathcal{M}_4$, leads to semi–finished (in–working) product.

The markings of places p_1 and p_2 represent the number of assembly servers for t_4 and t_3 respectively. The marking of places p_3, p_5, and p_9 represent the availability of parts to be processed (raw materials), while the marking of places p_4, p_6, p_7 and p_8 represent the availability of semi–finished products. Places p_{11} and p_{12} ensure that machines t_1 and t_2 work alternatively.

The Petri net model in Fig. 2 is a strongly connected event graph with $m = |P| = 12$ and $n = |T| = 7$. There exist eleven elementary circuits, that correspond to an equal number of P-invariants. If we assume that the initial marking of the net is that in Fig. 2, we have (here to avoid a heavy notation we denote as M_i the marking of place p_i)

$$\begin{cases} M_{11} + M_{12} = 1 & (1) \\ M_1 + M_3 + M_4 = 5 & (2) \\ M_1 + M_5 + M_6 = 5 & (3) \\ M_1 + M_3 + M_6 + M_{11} = 6 & (4) \\ M_1 + M_4 + M_5 + M_{12} = 5 & (5) \\ M_1 + M_3 + M_4 + M_{11} + M_{12} = 6 & (6) \\ M_2 + M_8 + M_9 = 6 & (7) \\ M_2 + M_3 + M_4 + M_7 + M_{10} = 6 & (8) \\ M_2 + M_5 + M_6 + M_7 + M_{10} = 6 & (9) \\ M_2 + M_3 + M_6 + M_7 + M_{10} + M_{11} = 7 & (10) \\ M_2 + M_4 + M_5 + M_7 + M_{10} + M_{12} = 6. & (11) \end{cases} \tag{3}$$

We assume that the above set of P-invariants coincides with the macromarking, thus $B_{w_0} = \vec{b} = [1\ 5\ 5\ 6\ 5\ 6\ 6\ 6\ 6\ 7\ 6]^T$.

Moreover, we assume that the controller must enforce two specifications:

$$\begin{cases} M_3 + M_5 \leq 3 & (a) \\ M_9 \leq 3. & (b) \end{cases} \tag{4}$$

The first specification requires that at most 3 raw parts may be simultaneously waiting to be processed by either machine $\mathcal{M}_1$ or $\mathcal{M}_2$. The second specification requires that at most 3 raw parts may be waiting to be processed by machine $\mathcal{M}_5$.

If the marking of the net is measurable, then the controlled net is live, being an event graph with generalized mutual exclusion constraints.

If the marking of the plant is not measurable, an observer must be used in the control loop. The resulting closed loop behaviour is represented in the reachability graph in Fig. 3. Here each node is labeled as: $(M \mid M_w / B_w)$.

We can immediately observe that the error estimate $M - M_w$ decreases as the length of the observed word increases. Nevertheless, after the firing of t_3 we reach a blocking condition. In fact, the controller prevents the firing of both transitions t_6 and t_7 even if their firing is perfectly legal. This is due to the fact that there exists at least one marking in $\mathcal{M}(t_1 t_4 t_3 \mid V, \vec{b})$ that would produce the violation of one of

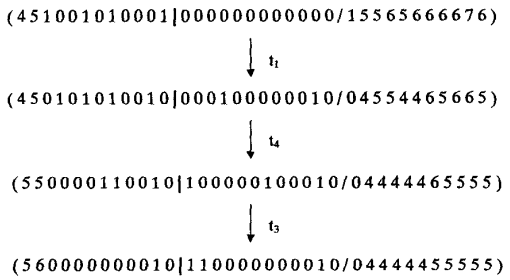

Fig. 3 Reachability graph of the net in Fig. 2 under control when no deadlock recovery procedure is applied.

the controller specifications if either transition t_6 or t_7 fires. In particular, the firing of t_6 may "potentially" violate the second specification (b), while the firing of t_7 does not guarantee the satisfaction of the first specification (a).

The first solution to this problem has been presented in [5] in the case of a particular manufacturing system and consists in the introduction of suitable time–out mechanisms. The limitation in the results in [5] is that the time–outs have been introduced with an "ad–hoc" reasoning.

In this paper we propose a general approach of deadlock recovery that is based on an original linear characterization of the set of deadlock markings.

V. A LINEAR ALGEBRAIC CHARACTERIZATION OF DEADLOCK MARKINGS

In this section we present one of the main contributions of the paper that consists in a linear algebraic characterization of deadlock markings that will be used in the next section to derive an efficient and systematic procedure for deadlock recovery. Such a characterization is valid for ordinary and structurally bounded Petri nets and is based on the following considerations and results.

We firstly observe that, by definition, the characteristic vector $\vec{s}$ of a siphon $\mathcal{S}$ is such that:

$$\forall t \in T \quad Post^T(\cdot,t)\cdot\vec{s} > 0 \;\Rightarrow\; Pre^T(\cdot,t)\cdot\vec{s} > 0. \quad (5)$$

Condition (5) means that if there exists a place p_i in the siphon $\mathcal{S}$ (i.e., $s_i = 1$) and t inputs in p_i (i.e., $Post(p_i,t) > 0$), then there must exist at least one place p_j in the siphon (i.e., $s_j = 1$) inputting in t (i.e., $Pre(p_j,t) > 0$).

Condition (5) can also be rewritten as a nonlinear inequality:

$$\forall t \in T \quad sign(Pre^T(\cdot,t)\cdot\vec{s}) \geq sign(Post^T(\cdot,t)\cdot\vec{s}) \quad (6)$$

where $sign(\vec{x})$ is a vector whose i–th component is 1 (resp., 0, -1) if the i–th component of $\vec{x}$ is positive (resp., null, negative).

Since the above inequality holds for all t, we can write

$$sign(Pre^T\cdot\vec{s}) \geq sign(Post^T\cdot\vec{s}). \quad (7)$$

We can finally state the following result.

Lemma 13 *A set of places $\mathcal{S} \subseteq P$ is a siphon of the net $N = (P,T,Pre,Post)$ if and only if its characteristic vector $\vec{s}$ is such that*

$$K_1 \cdot Pre^T \cdot \vec{s} \geq Post^T \cdot \vec{s}, \quad (8)$$

where $K_1 = \max_{t\in T} Post^T(\cdot,t)\cdot\vec{1}$.

Proof: We observe that $Post^T(\cdot,t)\cdot\vec{s} \leq Post^T(\cdot,t)\cdot\vec{1} \leq K_1$. Thus a vector $\vec{s} \in \{0,1\}^m$ is a solution of (8) if and only if it is a solution (7). ■

Secondly, let $M \in \mathbb{N}^m$ be a generic marking of N. If M corresponds to a reachable marking of the net such that the siphon $\mathcal{S}$ with characteristic vector $\vec{s}$ is empty, then

$$M^T \cdot \vec{s} = 0. \quad (9)$$

As in the previous case, we observe that in the case of structurally bounded Petri nets, equation (9) can be easily converted to a linear equivalent equation.

Lemma 14 *Given a structurally bounded net $N = (P,T,Pre,Post)$, a siphon $\mathcal{S}$ with characteristic vector $\vec{s}$ is empty at marking M if and only if*

$$K_2 \cdot \vec{s} + M \leq K_2 \cdot \vec{1}_m, \quad (10)$$

where K_2 is a sufficiently large positive integer. More precisely, K_2 should be chosen greater or equal to the maximum structural bound of p, for $p \in P$ [14], where structural bounds can be determined by using any LP software.

Proof: Equation (10) implies that if for a given j, $s_j = 0$ (i.e., place p_j does not belong to the siphon) then no constraint exists on the marking of p_j, since the equation $M(p_j) \leq K_2$ is satisfied for all reachable markings. On the contrary, if $s_j = 1$ (i.e., place p_j belongs to the siphon) then p_j must be empty. ■

An analogous linear characterization has been already proposed by Chu and Xie in [2].

Thirdly and finally, to completely characterize the set of deadlock markings we use the following results.

Lemma 15 ([13]) *Let N be a marked net. If $M \in \mathbb{N}^m$ is a dead marking then $\mathcal{S}_\emptyset = \{p \in P \mid M(p) = 0\}$ is an unmarked siphon of N.*

We restate the previous result in a slightly different form.

Lemma 16 *Let N be an ordinary marked net. A marking $M \in \mathbb{N}^m$ is a dead marking iff the two statements hold:*

(i) $\mathcal{S}_\emptyset = \{p \in P \mid M(p) = 0\}$ is an unmarked siphon of N;

(ii) $\forall t \in T$, ${}^\bullet t \cap \mathcal{S} \neq \emptyset$.

Proof: (if) It immediately follows from (ii) and the definition of enabled transitions. Condition (i) and (ii) together imply that every transition has at least one empty input place, thus no transition is enabled to fire.

(only if) Condition (i) follows from lemma 15 while condition (ii) follows from the fact that dead transitions must have in M at least an empty input place, that by definition belong to $\mathcal{S}_\emptyset$. ■

On the basis of the above lemmas, we can finally state the main result.

Theorem 17 *Given a structurally bounded net N with m places, a marking $M \in \mathbb{N}^m$ is a deadlock marking if and only if there exists a vector $\vec{s} \in \{0,1\}^m$ such that the following set of linear equations is satisfied:*

$$\mathcal{D}(Pre, Post) := \begin{cases} K_1 \cdot Pre^T \cdot \vec{s} \geq Post^T \cdot \vec{s} & (a) \\ K_2 \cdot \vec{s} + M \leq K_2 \cdot \vec{1}_m & (b) \\ \vec{s} + M \geq \vec{1}_m & (c) \\ Pre^T \cdot \vec{s} \geq \vec{1} & (d) \\ M \in \mathbb{N}^m & (e) \\ \vec{s} \in \{0,1\}^m & (f) \end{cases} \tag{11}$$

Proof: A marking M is dead if and only if there exists a siphon $\mathcal{S}_\emptyset$ as defined in Lemma 16.

The characteristic vector $\vec{s}$ must satisfy (a) by Lemma 13 and (b) by Lemma 14. Furthermore, by definition of $\mathcal{S}_\emptyset$, if a place p is empty at marking M then it must belong to the siphon, and this is imposed by (c). Finally, (d) states that for any transition t there exists at least one input place that is empty at M, and that consequently belongs to siphon $\mathcal{S}_\emptyset$. ■

In the following section we discuss in detail how this linear algebraic characterization may be successfully used to derive an efficient procedure for deadlock recovery.

VI. A SYSTEMATIC PROCEDURE FOR DEADLOCK RECOVERY

As discussed in detail in the previous sections, the controller may disable transitions whose firing is perfectly legal, and because of this it may be the case that the controlled system is blocking. Such an example has also been presented in subsection IV.A.

Now, let us propose a general approach to automatically recover from a blocking condition. The approach is essentially based on the linear algebraic characterization of deadlock markings given by equation (11). In particular, the above linear characterization is used to derive additional information on the actual marking of the net, so as to improve the marking estimate and restrict the set of $w-$consistent markings, where w is the observed word that lead N to a deadlock.

The structure of algorithm 12 that enables us to compute the control pattern γ keeps unaltered. The only variation involves the computation of $\overline{m}$. More precisely, we assume that, whenever the controlled system enters a blocking condition, and consequently no further event is observed for a sufficiently long time, then the deadlock recovery procedure is applied, that consists in recomputing the control pattern γ via algorithm 12, where $\overline{m}$ is evaluated as follows.

We consider the net obtained from the original one by simply removing all transitions whose firing is forbidden by the controller, and all arcs connected to them. In such a way we obtain a new net N' with m places and $n' < n$ transitions. We denote Pre' and $Post'$ the pre- and post- incidence matrices of N'. By virtue of the above linear characterization, we can be sure that if N' is in a blocking condition, then there exists at least one unmarked siphon $\mathcal{S}$, and the system is in a marking M such that $(M, \vec{s}) \in \mathcal{D}(Pre', Post')$. Obviously, the unknown marking M should also belong to the $w-$consistent set $\mathcal{M}(w\,|V, \vec{b})$. Therefore, when re–applying algorithm 12 we compute $\overline{m}$ as

$$\overline{m} := \max\{ \quad \vec{w_j}^T \cdot M \mid \exists\, \vec{s} \in \{0,1\}^m \text{ and } (M, \vec{s}) \in \mathcal{D}(Pre', Post') \cap \mathcal{M}(w\,|V, \vec{b})\}. \tag{12}$$

Once the control algorithm 12 has been applied with the updated definition of $\overline{m}$, three different cases may occur.

1. The *updated control pattern* — i.e., the set of transitions that are disabled by the controller computed using equation (12) — is less restrictive than the original control pattern and there exists at least one transition that may fire.

   In such a case a new transition t fires and the net recovers from the blocking condition. To continue applying the marking estimation algorithm, we suggest to use the linear algebraic characterization above to further improve the previous estimate μ_w. This simply requires the solution of m linear integer programming problems (LIPP), one for each place $p_i \in P$:

   $$\begin{cases} \min M(p_i) \\ s.t. \\ (M, \vec{s}) \in \mathcal{D}(Pre', Post') \\ M \in \mathcal{M}(w\,|V, \vec{b}) \end{cases} \tag{13}$$

   Note that the choice of the performance index originates from the requirement that the marking estimate should always be a lower bound of the actual marking.

   Now, let $\tilde{\mu}_w = [M^*(p_1) \cdots M^*(p_m)]^T$ be the new updated estimate, where $M^*(p_i)$ is the solution of the i–th LIPP. We can also redefine the set of $w-$consistent markings as

   $$\tilde{\mathcal{M}}(w \mid V, \vec{b}) = \{ \quad M \in \mathbb{N}^m \mid M \geq \tilde{\mu}_w, \; V^T \cdot M = V^T \cdot \tilde{\mu}_w + B_w\}, \tag{14}$$

   and proceed with the marking estimation procedure at step 4 of algorithm 3, where $t_{\alpha_i} = t$.

2. The updated control pattern is less restrictive than the original control pattern but the net is still deadlocked.

   In such a case the deadlock recovery procedure needs to be applied for a second time. Note that the net N' this second time is not the same as it was the first time, because the set of disabled transitions is different in the two cases.

3. The updated control pattern is as restrictive as the original control pattern and thus the net is still deadlocked.

   In such a case the deadlock recovery procedure reveals to be inefficient.

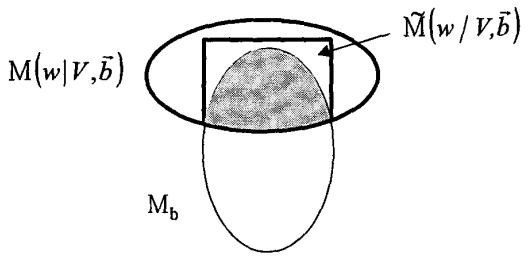

Fig. 4 Generic inclusion relationship among sets $\mathcal{M}(w \mid V,\vec{b})$, $\tilde{\mathcal{M}}(w \mid V,\vec{b})$ and $\mathcal{M}_b$.

So far, we have not been able to determine necessary and sufficient conditions to characterize those cases in which the deadlock recovery procedure works. This topic will be the object of our future research.

To conclude, we want to make an important remark concerning the procedure suggested in item 1 for the updating of the set of $w-$consistent markings. To this aim, let us consider Fig. 4. It sketches the generic inclusion relationship among sets $\mathcal{M}(w \mid V,\vec{b})$, $\tilde{\mathcal{M}}(w \mid V,\vec{b})$ and $\mathcal{M}_b$, where

$$\mathcal{M}_b = \{M \mid \exists\, \vec{s} \in \{0,1\}^m : (M,\vec{s}) \in \mathcal{D}(Pre', Post')\}$$

denotes the set of blocking markings induced in the closed loop system by the observer. If a blocking condition occurs, then the sets $\mathcal{M}(w \mid V,\vec{b})$ and $\mathcal{M}_b$ are not disjoint. The most precise solution in terms of marking estimation updating would be that of assuming the new set of $w-$consistent markings coincident with the dark area in Fig. 4, i.e.,

$$\begin{aligned}\tilde{\mathcal{M}}'(w \mid V,\vec{b}) \;&= \{M \in \mathbb{N}^m \mid \exists\, \vec{s} \in \{0,1\}^m,\\ &\quad (M,\vec{s}) \in \mathcal{D}(Pre', Post') \cap \mathcal{M}(w\,|V,\vec{b})\}.\end{aligned} \tag{15}$$

Nevertheless, this would require a significant increasing in the computational complexity of the procedure at the following steps. Thus, we retain that a more convenient solution is that of assuming $\tilde{\mathcal{M}}(w \mid V,\vec{b}) \subset \mathcal{M}(w \mid V,\vec{b})$ as defined in equation (14).

A. Numerical example

Let us consider again the assembly system in subsection IV.A. As discussed in detail before, the use of the marking estimate rather than the real marking of the net, may lead the closed loop system to a blocking condition.

In this section we show how the deadlock recovery procedure may be efficiently applied to the net in Fig. 2. If we assume that the initial marking is that in Fig. 2, then the first blocking condition occurs after the firing of the sequence $w = t_1t_4t_3$, as already discussed in subsection IV.A. The corresponding value of marking M_w, as well as that of the estimate μ_w and the bound B_w, may be seen in Fig. 3.

At this point, if we apply the updated version of algorithm 12, we obtain that the controller enables all transitions to fire, and among them there is one transition, t_6, that may actually fire (thus we are in case 1). Note that, in this case matrices Pre' and $Post'$ have been obtained by simply removing both the sixth and the seventh column from the pre- and post- incidence matrix of the original Petri net.

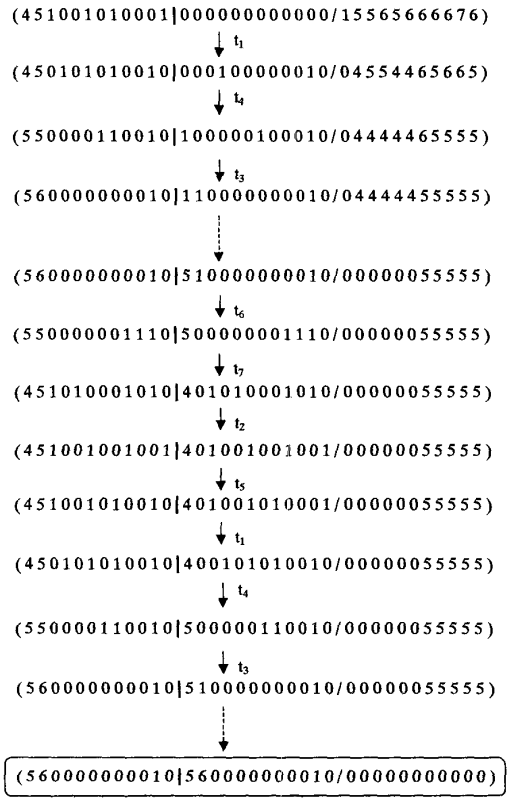

Fig. 5 One possible evolution of the net in Fig. 2 under control when the deadlock recovery procedure is applied.

Now, before computing the marking estimation after the firing of t_6, we solve twelve LIPP of the form (13) (one for each place) so as to compute the updated value of the marking estimate μ_w. Following the procedure previously illustrated, we determine $\tilde{\mu}_w = [5\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0]^T$, and the new value of the bound B_w becomes $\tilde{B}_w = [0\ 0\ 0\ 0\ 0\ 0\ 5\ 5\ 5\ 5\ 5\ 5]^T$.

To completely demonstrate the effectiveness of the proposed approach, in Fig. 5 we have reported one possible evolution of the closed loop system with observer, assuming that also the deadlock recovery procedure is applied. We used the same notation as that in Fig. 3, apart from the introduction of a dashed arrow denoting that no transition has fired, but only the marking estimation has been updated. As it can be seen, at the end of this evolution path, the marking is completely reconstructed and no further deadlock may occur.

The same can be repeated for any other evolution starting from the initial marking in Fig. 2, as it can be easily seen by looking at the whole reachability graph, that has not been reported here for brevity's requirements.

VII. CONCLUSIONS

We considered the state feedback control of Petri nets under the assumptions that the state is not measurable but can only be estimated, while the control objective is that of enforcing a set of GMECS. We showed by means of an example that the use of an estimate can easily lead to an "observer induced" deadlock and we have presented a general approach to recover from such a deadlock.

References

[1] 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.

[2] F. Chu, X. Xie, "Deadlock Analysis of Petri Nets Using Siphons and Mathematical Programming," *IEEE Trans. on Robotics and Automation*, Vol. 13, No. 6, pp. 793–804, 1997.

[3] Di Cesare, F., G. Harhalakis, J.M. Proth, M. Silva and F.B. Vernadat, *Practice of Petri nets in manufacturing*, Chapman and Hall, 1993.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] A. Giua, C. Seatzu, "Observability Properties of Petri Nets," *Proc. 39th Int. Conf. on Decision and Control*, Sydney, Australia, pp. 2676–2681, December 2000. A longer version has been conditionally accepted for publication on the *IEEE Trans. on Automatic Control*.

[9] 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.

[10] 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.

[11] 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.

[12] T. Murata, "Petri nets: properties, analysis and applications," *Proc. IEEE*, Vol. Proc. 77, N. 4, pp. 541–580, April 1989.

[13] W. Reisig, *Petri Nets. An Introduction*, Springer–Verlag, Berlin Heidelber New York Tokyo, 1982.

[14] M. Silva, J.M. Colom, "On the computation of structural synchronic invariants in P/T nets," in *Advances in Petri Nets 1988*, Springer–Verlag, New York, 1989.

[15] 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.

[16] 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.

[17] 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.