

Implicit Siphon Control and its Role in the Liveness Enforcing Supervision of Sequential Resource Allocation Systems

Spyros Reveliotis, *Senior Member, IEEE*

Abstract—The work presented in this paper seeks (i) to correct and generalize some previously published results regarding the siphon control in PT-ordinary Petri nets, and subsequently (ii) to employ this generalized framework in order to provide an alternative explanation of the way in which certain methodologies, that have been proposed in the past, enforce the liveness of a particular class of PT-ordinary Petri nets modelling resource allocation. The derived characterizations provide a *unifying* framework for analyzing and interpreting the aforementioned methodologies, and also they reveal that approaches that have been considered as disparate in the current literature, can actually be “*mixed*” towards the development of an ever richer set of liveness enforcing supervisory control policies for the considered class of Petri nets.

Index Terms—sequential resource allocation systems, system design and verification using Petri nets, deadlock resolution and liveness enforcing supervision

I. INTRODUCTION

The role and the significance of the structural object of siphon for the deadlock and liveness analysis of many Petri net (PN) classes are well-documented in the relevant literature, e.g., [1], [2], [3], [4], [5], [6], [7]. Generally speaking, the development of deadlock markings in any PN system, as well as the non-liveness of a considerable number of PN sub-classes, can be attributed to the insufficient marking of some of the system siphons. In the particular case that the considered net is PT-ordinary – i.e., the firing of any transition requires at most one token from each of the net places – the aforementioned insufficiently marked siphons are *empty* siphons. Siphons that never empty during the evolution of the net marking are said to be *controlled*. Hence, for PT-ordinary PN’s, ensuring that every siphon is controlled guarantees the net deadlock-freedom, and, in certain cases, the net liveness.

The work presented in this paper seeks (i) to correct and generalize some results regarding the siphon control in PT-ordinary Petri nets, originally published in [8], and subsequently (ii) to employ this generalized framework in order to provide an alternative explanation of the way in which certain methodologies, that have been proposed in the past, enforce the liveness of a particular class of PT-ordinary Petri nets modelling resource allocation. Beyond providing a novel *unifying* framework for analyzing the aforementioned methodologies, the presented results enable also their “*mixing*” towards the development of an ever richer set of liveness enforcing supervisory control policies for the considered class of Petri nets.

The rest of the paper is organized as follows: Section II provides a brief introduction to the Petri net modelling framework, presenting all the notation, concepts and results that are necessary for the subsequent developments. Section III presents our generalization of the results developed in [8], which takes the form of some new sufficiency tests for assessing whether any given siphon of a marked PN is controlled or not. The last part of this section discusses also some problems with some of the statements and derivations in [8], that were revealed by our new developments. Section IV first establishes the generalizing power of the results derived in Section III, by demonstrating the ability of this set of results to analyze and interpret the efficacy of a class of liveness enforcing supervisory control policies, that have been developed for certain PT-ordinary PN sub-classes where liveness is equivalent to the absence of uncontrolled siphons, and are not covered by the results presented in [8]. Subsequently, the last part of this section briefly surveys the liveness enforcing supervisory control policies studied in [8], and it discusses how the two classes of policies addressed in this section, can actually be “*mixed*” towards the development of an ever richer set of liveness enforcing supervisory control policies for the considered class of PT-ordinary Petri nets. Finally, Section V concludes the paper and suggests directions for future work.

II. PETRI NET PRELIMINARIES

Petri net Definition A formal definition of the Petri net model is as follows:

Definition 1: [9] A (*marked*) *Petri net (PN)* is defined by a quadruple $\mathcal{N} = (P, T, W, M_0)$, where

- P is the set of *places*,
- T is the set of *transitions*,
- $W : (P \times T) \cup (T \times P) \rightarrow Z_0^+$ is the *flow relation*,¹ and
- $M_0 : P \rightarrow Z_0^+$ is the net *initial marking*, assigning to each place $p \in P$, $M_0(p)$ *tokens*.

The first three items in Definition 1 essentially define a *weighted bipartite digraph* representing the system *structure* that governs its underlying dynamics. The last item defines the system *initial state*. A conventional graphical representation of the net structure and its marking depicts nodes corresponding to places by empty circles, nodes corresponding to transitions by bars, and the tokens located at the various places by small filled circles. The flow relation W is depicted by directed

¹In this work, Z_0^+ denotes the set of nonnegative integers, and Z^+ denotes the set of strictly positive integers.

edges that link every nodal pair for which the corresponding W -value is non-zero. These edges point from the first node of the corresponding pair to the second, and they are also labelled – or, “*weighed*” – by the corresponding W -value. By convention, absence of a label for any edge implies that the corresponding W -value is equal to unity.

PN structure-related concepts and properties Given a transition $t \in T$, the set of places p for which $(p, t) > 0$ (resp., $(t, p) > 0$) is known as the set of *input* (resp., *output*) places of t . Similarly, given a place $p \in P$, the set of transitions t for which $(t, p) > 0$ (resp., $(p, t) > 0$) is known as the set of *input* (resp., *output*) transitions of p . It is customary in the PN literature to denote the set of input (resp., output) transitions of a place p by $\bullet p$ (resp., $p\bullet$). Similarly, the set of input (resp., output) places of a transition t is denoted by $\bullet t$ (resp., $t\bullet$). This notation is also generalized to any set of places or transitions, X , e.g. $\bullet X = \bigcup_{x \in X} \bullet x$.

The ordered set $X = \langle x_1 \dots x_n \rangle \in (P \cup T)^*$ is a *path*, if and only if (iff) $x_{i+1} \in x_i\bullet, i = 1, \dots, n-1$. Furthermore, a path X is characterized as a *circuit* iff $x_1 \equiv x_n$.

A PN with a flow relation W mapping onto $\{0, 1\}$ is said to be *ordinary*. If only the restriction of W to $(P \times T)$ maps on $\{0, 1\}$, the PN is said to be *PT-ordinary*. An ordinary PN such that (s.t.) $\forall t \in T, |t\bullet| = |\bullet t| = 1$, is characterized as a *state machine*, while an ordinary PN s.t. $\forall p \in P, |p\bullet| = |\bullet p| = 1$, is characterized as a *marked graph*.

A PN is said to be *pure* if $\forall (x, y) \in (P \times T) \cup (T \times P), W(x, y) > 0 \Rightarrow W(y, x) = 0$. The flow relation of pure PN's can be represented by the *flow matrix* $\Theta = \Theta^+ - \Theta^-$ where $\Theta^+(p, t) = W(t, p)$ and $\Theta^-(p, t) = W(p, t)$.

PN dynamics-related concepts and properties In the PN modelling framework, the system state is represented by the net *marking* M , i.e., a function from P to Z_0^+ that assigns a *token* content to the various net places. The net marking M is initialized to marking M_0 , introduced in Definition 1, and it subsequently evolves through a set of rules summarized in the concept of *transition firing*. A concise characterization of this concept has as follows: Given a marking M , a transition t is *enabled* iff for every place $p \in \bullet t, M(p) \geq W(p, t)$, and this is denoted by $M[t]$. $t \in T$ is said to be *disabled* by a place $p \in \bullet t$ at M iff $M(p) < W(p, t)$. Furthermore, a place $p \in P$ for which there exists $t \in p\bullet$ s.t. $M(p) < W(p, t)$ is said to be a *disabling* place at M . Given a marking M , a transition t can be *fired* only if it is enabled in M , and firing such an enabled transition t results in a new marking M' , which is obtained from M by removing $W(p, t)$ tokens from each place $p \in \bullet t$, and placing $W(t, p')$ tokens in each place $p' \in t\bullet$. For pure PN's, the marking evolution incurred by the firing of a transition t can be concisely expressed by the *state equation*:

$$M' = M + \Theta \cdot \mathbf{1}_t \quad (1)$$

where $\mathbf{1}_t$ denotes the unit vector of dimensionality $|T|$ and with the unit element located at the component corresponding to transition t .

The set of markings reachable from the initial marking M_0 through any *fireable* sequence of transitions is denoted by $R(\mathcal{N}, M_0)$ and it is referred to as the net *reachability space*. In the case of pure PN's, a necessary condition for

$M \in R(\mathcal{N}, M_0)$ is that the following system of equations is feasible in z :

$$M = M_0 + \Theta z \quad (2)$$

$$z \in (Z_0^+)^{|T|} \quad (3)$$

A PN $\mathcal{N} = (P, T, W, M_0)$ is said to be *bounded* iff all markings $M \in R(\mathcal{N}, M_0)$ are bounded. \mathcal{N} is said to be *structurally bounded* iff it is bounded for any initial marking M_0 . \mathcal{N} is said to be *reversible* iff $M_0 \in R(\mathcal{N}, M)$, for all $M \in R(\mathcal{N}, M_0)$. A transition $t \in T$ is said to be *live* iff for all $M \in R(\mathcal{N}, M_0)$, there exists $M' \in R(\mathcal{N}, M)$ s.t. $M'[t]$; non-live transitions are said to be *dead* at those markings $M \in R(\mathcal{N}, M_0)$ for which there is no $M' \in R(\mathcal{N}, M)$ s.t. $M'[t]$. PN \mathcal{N} is *quasi-live* iff for all $t \in T$, there exists $M \in R(\mathcal{N}, M_0)$ s.t. $M[t]$; it is *weakly live* iff for all $M \in R(\mathcal{N}, M_0)$, there exists $t \in T$ s.t. $M[t]$; and it is *live* iff for all $t \in T, t$ is live. A marking $M \in R(\mathcal{N}, M_0)$ is a (total) *deadlock* iff every $t \in T$ is dead at M .

Siphons and their role in the interpretation of the PN deadlock A *siphon* is a set of places $S \subseteq P$ such that $\bullet S \subseteq S\bullet$. A siphon S is *minimal* iff there exists no other siphon S' s.t. $S' \subset S$. A siphon S is said to be *empty* at marking M iff $M(S) \equiv \sum_{p \in S} M(p) = 0$. S is said to be *deadly marked* at marking M , iff every transition $t \in \bullet S$ is disabled by some place $p \in S$. Clearly, empty siphons are deadly marked siphons. It is easy to see that, if S is a deadly marked siphon at some marking M , then (i) $\forall t \in \bullet S, t$ is a dead transition in M , and (ii) $\forall M' \in R(\mathcal{N}, M), S$ is deadly marked. The next theorem connects total deadlocks arising in PN's to deadly marked siphons.

Theorem 1: [10] Given a deadlock marking M of a PN $\mathcal{N} = (P, T, W, M_0)$, the set of disabling places $S \subseteq P$ in M constitutes a deadly marked siphon.

In PT-ordinary PN's, disabling places are empty places. Hence, an immediate corollary of Theorem 1 is as follows:

Corollary 1: Given a deadlock marking M of a PT-ordinary PN $\mathcal{N} = (P, T, W, M_0)$, the set of disabling places $S \subseteq P$ in M constitutes an empty siphon.

Finally, as it was mentioned in the Introduction, a siphon S is said to be *controlled*, iff it never empties during the evolution of the net marking, i.e., $M(S) > 0, \forall M \in R(\mathcal{N}, M_0)$.

PN semiflows PN semiflows provide an analytical characterization of various concepts of *invariance* underlying the net dynamics. Generally, there are two types, p and t-semiflows, with a *p-semiflow* formally defined as a $|P|$ -dimensional vector y satisfying $y^T \Theta = 0$ and $y \geq 0$, and a *t-semiflow* formally defined as a $|T|$ -dimensional vector x satisfying $\Theta x = 0$ and $x \geq 0$. In the light of Equation 2, the invariance property expressed by a p-semiflow y is that $y^T M = y^T M_0$, for all $M \in R(\mathcal{N}, M_0)$. Similarly, Equation 2 implies that for any t-semiflow $x, M = M_0 + \Theta x = M_0$.

Given a p-semiflow y (resp., t-semiflow x) its *support* is defined as $\|y\| = \{p \in P \mid y(p) > 0\}$ (resp., $\|x\| = \{t \in T \mid x(t) > 0\}$). A p-semiflow y (resp., t-semiflow x) is said to be *minimal* iff there is no p-semiflow y' (resp., t-semiflow x') s.t. $\|y'\| \subset \|y\|$ (resp., $\|x'\| \subset \|x\|$).

PN merging We conclude our general discussion on the PN concepts and properties to be employed in the subsequent

parts of this work, by introducing a merging operation of two PN's: Given two PN's $\mathcal{N}_1 = (P_1, T_1, W_1, M_{01})$ and $\mathcal{N}_2 = (P_2, T_2, W_2, M_{02})$ with $T_1 \cap T_2 = \emptyset$ and $P_1 \cap P_2 = Q \neq \emptyset$ s.t. for all $p \in Q$, $M_{01}(p) = M_{02}(p)$, the PN \mathcal{N} resulting from the merging of the nets \mathcal{N}_1 and \mathcal{N}_2 through the place set Q , is defined by $\mathcal{N} = (P_1 \cup P_2, T_1 \cup T_2, W_1 \cup W_2, M_0)$ with $M_0(p) = M_{01}(p)$, $\forall p \in P_1 \setminus P_2$; $M_0(p) = M_{02}(p)$, $\forall p \in P_2 \setminus P_1$; $M_0(p) = M_{01}(p) = M_{02}(p)$, $\forall p \in P_1 \cap P_2$.

III. IMPLICIT SIPHON CONTROL

In this section we develop some new tests for identifying controlled siphons in any given marked PN \mathcal{N} . As it was mentioned in the introductory section, the results derived in this section generalize, but also correct, some results initially developed in [8]. The following two concepts are instrumental for the development of the proposed tests:

Definition 2: Consider a marked PN $\mathcal{N} = (P, T, W, M_0)$ and a vector $v \in \mathfrak{R}^{|P|}$, where \mathfrak{R} denotes the set of reals. Then, for any marking $M \in R(\mathcal{N}, M_0)$, the *generalized compound marking* generated by v , is defined by

$$GCM(M, v) = \sum_{p \in P} v(p)M(p) = v^T M \quad (4)$$

The vector v will be called the *generator* of $GCM(M, v)$.

Notice that in the particular case that $v(p) \in \{0, 1\}$, $\forall p \in P$, a $GCM(M, v)$ reduces to the *compound marking* of the place subset P^v defined by the support of v . In the following, P^v will denote more generally the set of places corresponding to non-zero elements of v .

Definition 3: Consider a *pure* marked PN $\mathcal{N} = (P, T, W, M_0)$ and a GCM generator $v \in \mathfrak{R}^{|P|}$. Then, the *net flow (vector)* of v is defined by

$$NF(v) = v^T \Theta \quad (5)$$

where Θ denotes the flow matrix of \mathcal{N} .

Notice that $NF(v)$ is a $|T|$ -dimensional row vector. Furthermore, in the light of Equation 1, the components of $NF(v)$ have the following very intuitive interpretation: For every transition $t \in T$, $NF(v; t)$ denotes the net change of $GCM(M, v)$ resulting by the firing of transition t at M .

The next definition connects the GCM and NF concepts to the concept of siphon.

Definition 4: Consider a siphon S of a *pure* marked PN $\mathcal{N} = (P, T, W, M_0)$. The *characteristic vector*² of S is a $|P|$ -dimensional binary vector λ_S such that

$$\forall p \in P, \quad \lambda_S(p) = 1 \iff p \in S \quad (6)$$

The characteristic vector λ_S , of any given siphon S , can be considered as a GCM generator with $GCM(M, \lambda_S)$ being equal to the token content of siphon S at marking M . Furthermore, the components of the corresponding net flow vector $NF(\lambda_S)$ express the net change incurred to the siphon marking by the firing of any single transition $t \in T$.

²In order to facilitate the interpretation of the results presented in [8] in the context of the results presented in this work, we notice that this concept corresponds to the *characteristic P-vector* defined in [8].

Now we have all the necessary concepts in place in order to state and prove the main result of this section; this is done in the following theorem:

Theorem 2: Let S denote a siphon of a *pure* marked PN $\mathcal{N} = (P, T, W, M_0)$ such that

$$NF(\lambda_S) = \sum_{i=1}^n a_i NF(v^i) \quad (7)$$

where v^i , $i = 1, \dots, n$, are GCM generators of \mathcal{N} , and $a_i \in \mathfrak{R}$, $\forall i$. Then,

$$S \text{ is controlled in } \mathcal{N} \iff \lambda_S^T M_0 + G^* > 0 \quad (8)$$

where

$$G^* = \min_{M \in R(\mathcal{N}, M_0)} (M - M_0)^T \sum_{i=1}^n a_i v^i \quad (9)$$

Proof: Consider a marking $M \in R(\mathcal{N}, M_0)$. Then, there exists a vector $z \in (Z_0^+)^{|T|}$ such that $M = M_0 + \Theta z$ (c.f. Equations 2 and 3). Therefore,

$$\begin{aligned} M(S) &= \sum_{p \in S} M(p) \\ &= \lambda_S^T M \\ &= \lambda_S^T M_0 + \lambda_S^T \Theta z \\ &= \lambda_S^T M_0 + NF(\lambda_S) z \\ &= \lambda_S^T M_0 + \left[\sum_i a_i NF(v^i) \right] z \\ &= \lambda_S^T M_0 + \left[\sum_i a_i (v^i)^T \Theta \right] z \\ &= \lambda_S^T M_0 + \left[\sum_i a_i v^i \right]^T \Theta z \\ &= \lambda_S^T M_0 + \left[\sum_i a_i v^i \right]^T (M - M_0) \\ &= \lambda_S^T M_0 + (M - M_0)^T \sum_i a_i v^i \end{aligned} \quad (10)$$

Clearly, the right-hand-side of Equation 10 is minimized over $R(\mathcal{N}, M_0)$ by G^* , and therefore, S will be controlled if and only if the criterion of Equation 8 holds. ■

A siphon S controlled by means of the criterion of Theorem 2 will be characterized as an *implicitly* controlled siphon. The corresponding generator vectors v^i , $i = 1, \dots, n$, of Equation 7, will be called the *controlling generators* of S . In practice, the application of the criterion of Theorem 2 on any siphon S with respect to any given set of generator vectors $\{v^i : i = 1, \dots, n\}$ that satisfy Equation 7, is complicated by the fact that the constraint $M \in R(\mathcal{N}, M_0)$ cannot be represented easily – i.e., polynomially – by a set of linear constraints. Yet, one can compromise for a *sufficiency* test by relaxing the requirement $M \in R(\mathcal{N}, M_0)$ in Equation 9 to that expressed by the state Equations 2 and 3. We state the resulting criterion as a corollary.

Corollary 2: Let S denote a siphon of a *pure* marked PN $\mathcal{N} = (P, T, W, M_0)$ such that

$$NF(\lambda_S) = \sum_{i=1}^n a_i NF(v^i) \quad (11)$$

where v^i , $i = 1, \dots, n$, are *GCM* generators of \mathcal{N} , and $a_i \in \mathbb{R}$, $\forall i$. Also, let

$$G' = \min_{(M,z)} (M - M_0)^T \sum_{i=1}^n a_i v^i \quad (12)$$

s.t.

$$M = M_0 + \Theta z \quad (13)$$

$$M \geq 0, \quad z \in (Z_0^+)^{|T|} \quad (14)$$

Then,

$$\lambda_S^T M_0 + G' > 0 \implies S \text{ is controlled in } \mathcal{N} \quad (15)$$

Notice that the mathematical programming (MP) formulation involved in the criterion of Corollary 2 is a Mixed Integer Program (MIP), and therefore, it can be easily addressed through commercial solvers (c.f. [11], for instance)³. Next we present another criterion that is weaker than the criterion of Corollary 2, but it reveals the connection of the presented results to those derived in [8]. Furthermore, this new criterion can be simpler, from a computational standpoint.

Corollary 3: Let S denote a siphon of a pure marked PN $\mathcal{N} = (P, T, W, M_0)$ such that

$$NF(\lambda_S) = \sum_{i=1}^n a_i NF(v^i) \quad (16)$$

where v^i , $i = 1, \dots, n$, are *GCM* generators of \mathcal{N} , and $a_i \in \mathbb{R}$, $\forall i$. Also, for every $i \in \{1, \dots, n\}$, let $\underline{GCM}(v^i)$ and $\overline{GCM}(v^i)$ respectively denote a lower and an upper bound of $GCM(M, v^i)$, for all M such that

$$M = M_0 + \Theta z \quad (17)$$

$$M \geq 0, \quad z \in (Z_0^+)^{|T|} \quad (18)$$

Finally, let

$$G'' = \sum_{i:a_i>0} a_i [\underline{GCM}(v^i) - GCM(M_0, v^i)] + \sum_{i:a_i<0} a_i [\overline{GCM}(v^i) - GCM(M_0, v^i)] \quad (19)$$

Then,

$$\lambda_S^T M_0 + G'' > 0 \implies S \text{ is controlled in } \mathcal{N} \quad (20)$$

Proof: Notice that

$$\begin{aligned} (M - M_0)^T \sum_i a_i v^i &= \sum_i a_i (M^T v^i - M_0^T v^i) \\ &= \sum_i a_i [GCM(M, v^i) - \\ &\quad GCM(M_0, v^i)] \end{aligned} \quad (21)$$

³In fact, the integrality requirement for z can be further relaxed to $z \geq 0$, providing a test that is computationally easier, but also with diminished resolution power, compared to the test of Corollary 2.

The definitions of $\underline{GCM}(v^i)$ and $\overline{GCM}(v^i)$, when combined with Equations 12-14, 17-19 and 21, imply that

$$G' \geq G'' \quad (22)$$

But then, the validity of Corollary 3 follows from Corollary 2. \blacksquare

Beyond providing a sufficiency test for assessing whether a given siphon S is implicitly controlled by a set of *GCM* generator vectors $\{v^i : i = 1, \dots, n\}$, the result of Corollary 3 can also provide the basis for deploying a control mechanism that will actively enforce the implicit control of siphon S by some generator set $\{v^i : i = 1, \dots, n\}$. Under this approach, the upper and lower bounds $\overline{GCM}(v^i)$ and $\underline{GCM}(v^i)$, $i = 1, \dots, n$, are “*design parameters*”, and their values are chosen such that they guarantee the condition of Equation 20. The selected bounds can be subsequently enforced on the behavior of the original net by the addition of appropriate “*monitor places*”, according to the theory developed in [12], [13]. The following result, established in [8], strengthens further the viability of such a control scheme, as it implies that the entire set of siphons, \mathcal{S} , of a pure marked PN $\mathcal{N} = (P, T, W, M_0)$, can be potentially controlled by a set of generators, and corresponding control places, that are polynomially – in fact, linearly – related to the size of the net \mathcal{N} , where the latter is expressed by $|P| + |T|$; the implications of this possibility are further explored in the next section.

Proposition 1: [8] Given a pure marked PN $\mathcal{N} = (P, T, W, M_0)$, the rank of the space of net flow vectors $NF(\lambda_S)$, corresponding to the net siphons S , is bounded from above by $\min\{|P|, |T|\}$.

We conclude this section by noticing that the result of Corollary 3 subsumes the result of Theorem 1 in [8]. On the other hand, Theorems 2 and Corollary 1 of [8] are erroneous, because they fail to recognize properly the impact of the second term in the right-hand-side of Equation 19. More specifically, the analysis of [8] restricts the potential set of *GCM* generators, $\{v^i\}$, to the set of the siphon characteristic vectors, $\{\lambda_S\}$, and then it makes the erroneous assumption that, for any given PN $\mathcal{N} = (P, T, W, M_0)$ and siphon S , $M_0(S) = GCM(M_0, \lambda_S) \geq GCM(M, \lambda_S) = M(S)$, $\forall M \in R(\mathcal{N}, M_0)$. Clearly, if this assumption were true, the second term in the right-hand-side of Equation 19 would be identically zero for the considered set of *GCM* generators; hence, [8] systematically ignores this term in its derivations and the presented results. But, while it happens that the aforesaid assumption is satisfied by the siphons that constitute the main focus of attention in [8],⁴ the next example establishes that this assumption is not generally true, and therefore, the disputed term can have a significant impact even when the considered set of *GCM* generators is restricted to the set $\{\lambda_S\}$.

Example 1 Consider the marked PN $\mathcal{N} = (P, T, W, M_0)$ depicted in Figure 1. This net was shown in [5] to be live and reversible. A minimal siphon of \mathcal{N} is $S = \{p_{13}, p_{23}, r_1, r_2, r_3\}$, with $M_0(S) = 4$. Next, consider the marking M of \mathcal{N} with $M(p_{10}) = M(p_{11}) = M(p_{12}) = M(p_{13}) = M(r_2) = 1$, $M(p_{20}) = 4$, and $M(p) = 0$ for

⁴and therefore, the results derived in Section VI of that paper remain correct

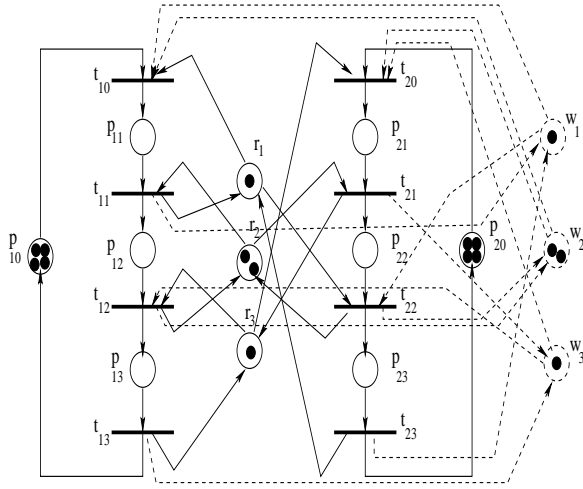


Fig. 1. The Petri net considered in Example 1

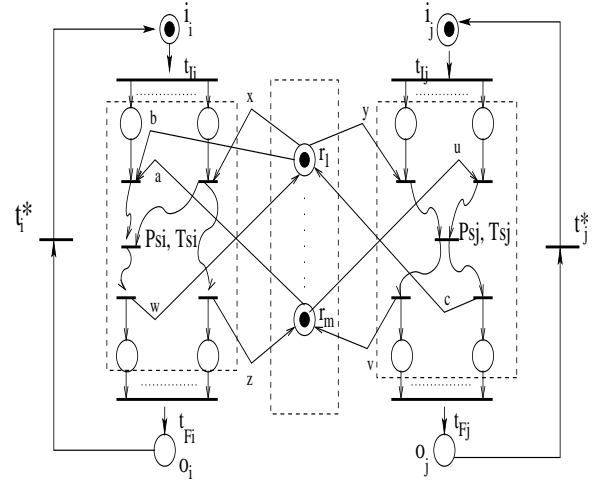


Fig. 2. The generic structure of *process-resource* nets

every other $p \in P$. It can be easily checked that $M = M_0[t_{10}t_{11}t_{12}t_{10}t_{11}t_{10} >$, and therefore, $M \in R(\mathcal{N}, M_0)$. Furthermore, notice that $M(S) = 2$. Finally, since \mathcal{N} is live and reversible, $M_0 \in R(\mathcal{N}, M)$. But then, the marked PN $\mathcal{N}' = (P, T, W, M'_0 \equiv M)$ constitutes a marked PN that contains a siphon S for which $\max_{M \in R(\mathcal{N}, M'_0)} M(S) \geq M_0(S) > M'_0(S)$, and establishes the fallacy of the aforementioned assumption of [8]. \diamond

In the next section it is shown that, while the siphon control criteria of Corollaries 2 and 3 encompass all the relevant results developed in [8], they can also support the analysis and interpret the correctness of an additional set of liveness enforcing supervisory control policies, that have been developed for certain PT-ordinary PN sub-classes where liveness is equivalent to the absence of uncontrolled siphons, and are not covered by the results developed in [8]. Therefore, it can be claimed that the results developed in this section constitute a substantial generalization of the corresponding set of results developed in [8].

IV. THE ROLE OF IMPLICIT SIPHON CONTROL IN ASSESSING AND ENFORCING THE LIVENESS OF SEQUENTIAL RESOURCE ALLOCATION SYSTEMS

Process-resource nets and their ES^3PR sub-class In this section, we shift attention to a particular PN sub-class known as *process-resource* nets. Process-resource nets have been extensively used in the literature for modelling the contest of concurrently executing processes for a finite set of reusable resources; some general characterizations of these nets and extensive studies of their properties can be found in [4], [7], [10]. Generally speaking, these nets are obtained by merging a set of sub-nets modelling the sequential logic and the resource allocation associated with the execution of their process types, through the places modelling the availability of the shared resources. The resulting net structure is depicted in Figure 2: Tokens contained in place i_i (resp., i_j) represent process instances of type JT_i (resp., JT_j) waiting to initiate execution, while tokens in place o_i (resp., o_j) correspond to completed processes. The firing of transition t_{I_i} (resp., t_{I_j})

models the initiation of a new process instance of type JT_i (resp., JT_j). Similarly, the firing of transition t_{F_i} (resp., t_{F_j}) models the completion of a process instance of type JT_i (resp., JT_j). The part of the net between transitions t_{I_i} and t_{F_i} (resp., t_{I_j} and t_{F_j}) encodes the sequential logic applying to the execution of any process instance of type JT_i (resp., JT_j). Places r_l are “monitor” places modelling the availability of the various resource types during the system operation. Their connectivity to the rest of the network encodes the resource requests posed by the various processing stages. Finally, transitions t_i^* and t_j^* allow the repetitive execution of the logic encoded by the corresponding process sub-nets, by enabling the “re-circulation” of the tokens modelling the relevant process instances. It is customary in the literature to “collapse” the path $\langle o_i, t_i^*, i_i \rangle$ to a single place p_{o_i} that is known as the “idle (state) place” of the corresponding process sub-net.

A development that has been extremely useful in the past studies of process-resource nets, is their classification in a taxonomy based on (i) the specific structure of the involved process sub-nets, and (ii) the structure of the restriction of the net flow relation, W , on $(P_R \times T) \cup (T \times P_R)$, where P_R denotes the set of resource places r_l . Following this practice, in the following we focus on a particular class of process-resource nets in which (i) the sequential logic governing the execution of the various process types is characterized by *acyclic state machines*, and (ii) $W(r_l, t) \in \{0, 1\}$, $\forall r_l, t$. The first of the above restrictions implies that the considered sub-class allows for choice – or “routing flexibility”, in the relevant terminology – but it also requires that any activated process remains a single atomic entity during its sojourn into the system (i.e., no task parallelization is allowed). The second restriction requires that resources from any particular type can be acquired by a process only one unit at a time (however, they can be accumulated and released in larger quantities). The resulting process-resource net sub-class belongs to the class of PT-ordinary PN’s and it is known as the class of *Extended Simple Sequential Systems of Processes with Resources*, or more briefly, as *ES³PR nets* [14]. A formal definition of

ES^3PR nets is as follows:

Definition 5: [14] Let $I_N = \{1, 2, \dots, m\}$ be a finite set of indices. A (well-marked) *Extended Simple Sequential System of Processes with Resources* – or, more briefly, an ES^3PR net – is a PT-ordinary marked PN $\mathcal{N} = (P, T, W, M_0)$, where:

- 1) $P = P_S \cup P_0 \cup P_R$ is a partition such that
 - a) $P_S = \cup_{i \in I_N} P_{S_i}$, with $P_{S_i} \cap P_{S_j} = \emptyset$ for all $i \neq j$.
 - b) $P_0 = \cup_{i \in I_N} \{p_{0_i}\}$ (idle state places).
 - c) $P_R = \{r_1, r_2, \dots, r_n\}$, $n > 0$.
- 2) $\forall i \in I_N$, the sub-net \mathcal{N}_i generated by $P_{S_i} \cup \{p_{0_i}\}$ and the transition subset T_i connected to these places, is a strongly connected state machine, such that every cycle contains place p_{0_i} .
- 3) $\forall r_l \in P_R$, \exists a unique minimal p -semiflow Y_{r_l} such that $\{r_l\} = \|Y_{r_l}\| \cap P_R$, $P_0 \cap \|Y_{r_l}\| = \emptyset$, $P_S \cap \|Y_{r_l}\| \neq \emptyset$ and $Y_{r_l}(r_l) = 1$.
- 4) $\forall r_l \in P_R$, $\forall t \in r_l^\bullet$, $W(r_l, t) = 1$.
- 5) $P_S = \cup_{r_l \in P_R} (\|Y_{r_l}\| \setminus \{r_l\})$.
- 6) \mathcal{N} is a connected net.
- 7) The initial marking M_0 satisfies the following conditions:
 - a) $\forall p \in P_S$, $M_0(p) = 0$.
 - b) $\forall p \in P_0$, $M_0(p) > 0$.
 - c) $\forall r_l \in P_R$, $M_0(r_l) \geq \max_{p \in \|Y_{r_l}\|} \{Y_{r_l}(p)\}$.

Item 3 of Definition 5 expresses the conservative – or *reusable* – nature of the system resources. Furthermore, the combination of items 2, 3 and 7 implies that every transition sequence σ defined by a circuit of some sub-net \mathcal{N}_i , $i \in I_N$, that leads from idle place p_{0_i} back to it, is fireable from M_0 , and therefore, \mathcal{N} is *quasi-live*.

A desirable property for process-resource nets is that they are *reversible*, since this property implies that all active processes in the underlying system can always proceed to completion, no matter what is the running loading pattern. It has been shown [4], [7] that, for a large class of process-resource nets, the net reversibility is a concept equivalent to the net *liveness*. In addition, the possession of these two properties by a process-resource net of this class, is contingent upon the absence of certain types of insufficiently marked siphons from its reachability space. For the particular case of ES^3PR nets, the following result is a direct implication of Corollary 3 and Theorem 2 in [7]:

Theorem 3: A well-marked ES^3PR net $\mathcal{N} = (P_S \cup P_0 \cup P_R, T, W, M_0)$ is live and reversible *iff* all its siphons are controlled.

Algebraic Liveness-Enforcing Supervisors (LES) for ES^3PR nets In general, the condition of Theorem 3 will not be immediately satisfied by any given well-marked ES^3PR net. However, it has been shown in the literature [10] that it is possible to establish the liveness and reversibility of these nets by imposing on their operation an additional set of linear inequalities to be observed by the net marking. One particular type of these inequalities seeks to constrain the number of process instances that are simultaneously executing certain subsets of processing stages. In the ES^3PR modelling framework, these inequalities take the form:

$$A \cdot M_S \leq b \quad (23)$$

where matrix A is an *incidence* – i.e., binary – matrix, M_S restricts the PN marking M to its components corresponding to places $p \in P_S$, and $b_l \in Z^+$, $\forall l$. The constraints expressed by Equation 23 are subsequently enforced on the considered ES^3PR net, by augmenting it with a controlling sub-net that is readily constructed through the theory of *control-place invariants* of [12], [13]. According to [13], each of the inequalities

$$A_{[l, \cdot]} \cdot M_S \leq b_l \quad (24)$$

can be imposed on the net behavior by superimposing on the original net structure a “*control place*” w_l , connected to the rest of the network according to the incidence vector

$$\theta_{w_l} = -A_{[l, \cdot]} \cdot \Theta_S \quad (25)$$

where Θ_S denotes the flow sub-matrix of the uncontrolled network $\mathcal{N} = (P_S \cup P_0 \cup P_R, T, W, M_0)$ corresponding to places $p \in P_S$. The initial marking of place w_l must be set to

$$M_0(w_l) = b_l \quad (26)$$

and the resulting controller imposes Constraint 24 on the system behavior by establishing the place invariant

$$A_{[l, \cdot]} \cdot M_S + M(w_l) = b_l \quad (27)$$

Let $P_W \equiv \cup_l \{w_l\}$. Equation 27, when interpreted in the light of item 3 of Definition 5, implies that the control places w_l , implementing each of the constraints in the LES-defining Equation 23, essentially play the role of *fictional* new resources in the dynamics of the net \mathcal{N}_c , that models the controlled system behavior. Hence, the controlled net \mathcal{N}_c remains in the broader class of process-resource nets. Next we show that the net \mathcal{N}_c is also an ES^3PR net.

Proposition 2: Consider the net \mathcal{N}_c obtained by enforcing on an ES^3PR net, $\mathcal{N} = (P_S \cup P_0 \cup P_R, T, W, M_0)$, a set of inequality constraints of the type expressed by Equation 23, according to the theory of control-place invariants. Then, under the assumption that $|w_l^\bullet| > 0, \forall l$, the net \mathcal{N}_c is also ES^3PR .

Proof: The conditions expressed by items 1, 2, 5, (7a) and (7b) of Definition 5 are immediately satisfied by the fact that the original net \mathcal{N} is ES^3PR . The condition of item 3 is met by Equation 27, while the condition of item (7c) results from Equation 27 and the facts that A is a binary matrix and $b_l \in Z^+$, $\forall l$. The condition of item 4 is satisfied by Equation 25 and the facts that A is a binary matrix and each process sub-net \mathcal{N}_i , $i \in I_N$, of \mathcal{N} is a state machine (c.f. item 2 of Definition 5). Finally, the condition expressed by item 6 of Definition 5 is satisfied by the fact that the original net \mathcal{N} is connected, and the posed assumption that $|w_l^\bullet| > 0, \forall l$. ■

Notice that the assumption $|w_l^\bullet| > 0, \forall l$, must be naturally satisfied by any algebraic LES of the type expressed by Equation 23 that contains no redundant constraints; such an LES will be characterized as *well-structured*. Then, the following corollary is an immediate implication of Theorem 3 and Proposition 2:

Corollary 4: Consider the net \mathcal{N}_c that is obtained by enforcing on an ES^3PR net, $\mathcal{N} = (P_S \cup P_0 \cup P_R, T, W, M_0)$, a well-structured algebraic LES of the type expressed by Equation 23, according to the theory of control-place invariants. \mathcal{N}_c

TABLE I
EXAMPLE 2: THE MINIMAL SIPHONS OF THE CONTROLLED NET \mathcal{N}_c

siphon	p_{10}	p_{11}	p_{12}	p_{13}	p_{20}	p_{21}	p_{22}	p_{23}	r_1	r_2	r_3	w_1	w_2	w_3
S_1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
S_2	0	0	0	0	1	1	1	1	0	0	0	0	0	0
S_3	0	1	0	0	0	0	0	1	1	0	0	0	0	0
S_4	0	0	1	0	0	0	1	0	0	1	0	0	0	0
S_5	0	0	0	1	0	1	0	0	0	0	1	0	0	0
S_6	0	1	0	0	0	1	1	1	0	0	0	1	0	0
S_7	0	1	1	0	0	1	1	0	0	0	0	0	1	0
S_8	0	1	1	1	0	1	0	0	0	0	0	0	0	1
S_9	0	0	1	0	0	0	0	1	1	1	0	0	0	0
S_{10}	0	0	0	1	0	0	1	0	0	1	1	0	0	0
S_{11}	0	0	0	1	0	0	0	1	1	1	1	0	0	0

TABLE II
EXAMPLE 2: THE GCM GENERATORS, v^i , EMPLOYED FOR THE EXPANSION OF THE NET FLOW VECTORS $NF(\lambda_{S_k})$, $k = 9, 10, 11$, AND THE ASSOCIATED BOUNDS USED IN THE EVALUATION OF THE CRITERION OF COROLLARY 3

generator	p_{10}	p_{11}	p_{12}	p_{13}	p_{20}	p_{21}	p_{22}	p_{23}	r_1	r_2	r_3	w_1	w_2	w_3	$\overline{GCM}(v^i)$	$\overline{GCM}(v^i)$
v^1	0	1	0	0	0	1	1	1	0	0	0	0	0	0	0	1
v^2	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	2
v^3	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	1
v^4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
v^5	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
v^6	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1

TABLE III
EXAMPLE 2: THE COORDINATES FOR THE EXPANSIONS OF $NF(\lambda_{S_k})$, $k = 9, 10, 11$, AS LINEAR COMBINATIONS OF $NF(v^l)$, $l = 1, \dots, 6$, AND THE OBTAINED VALUES FOR THE TEST OF COROLLARY 3

siphon	a_1	a_2	a_3	a_4	a_5	a_6	$\lambda_S^T M_0 + G''(S)$
S_9	0.0	-1.0	0.0	0.0	1.0	1.0	3-2=1
S_{10}	0.0	0.0	0.0	0.0	-1.0	-1.0	3-2=1
S_{11}	0.0	-1.0	0.0	0.0	0.0	0.0	4-2=2

the right-hand-side of Equation 19 is identically zero, and the only term that essentially defines the value of $G''(S)$ is the second term in that sum. \diamond

The above example renders clear the mechanism through which the GCM generator concept, introduced in Section III, explains the correctness of LES like RUN, that do not fall into the LES scope covered in [8]. Furthermore, the same example also reveals that, in this generalized regime, the second term in the right-hand-side of Equation 19, and the pertinent selection of the upper bounds involved in this term, can have an impacting role in the establishment of deadlock-freedom and/or the liveness and reversibility of ES^3PR nets through implicit siphon control. In the remaining part of this section, we briefly review the key results presented in [8] regarding the liveness enforcing supervision through implicit siphon control for a sub-class of the ES^3PR nets, known as S^3PR nets; this discussion will outline the connection of those past results to the results of Section III, and it will also reveal that it is possible to “mix” the control logic underlying the methodology of [8] with the control logic underlying the

algebraic LES of the type expressed by Equation 23, in order to obtain an even broader class of LES for the considered class of process-resource nets.

Liveness Enforcing Supervision of S^3PR nets through explicit control of “elementary” siphons The class of S^3PR nets constitutes a sub-class of ES^3PR nets, which is obtained by further stipulating that (i) every place $p \in P_S$ belongs to the support of only one of the p -semiflows Y_{r_l} introduced in item 3 of Definition 5, and (ii) $Y_{r_l}(p) \in \{0, 1\}$ for all $p \in P_S$ and $r_l \in P_R$.⁵ In order to develop an LES for S^3PR nets, [8] starts with the observation that, according to Theorem 3, the liveness and reversibility of any given S^3PR net, $\mathcal{N} = (P_S \cup P_0 \cup P_R, T, W, M_0)$, could be possibly enforced by adding a “monitor” place for every uncontrolled minimal siphon of \mathcal{N} , $S \in \mathcal{S}^{MU}$, that imposes the inequality:

$$\lambda_S^T M \geq \xi_S \quad (30)$$

⁵A more intuitive interpretation of these two requirements is that every processing stage requires only *one* unit from a *single* resource type for its execution.

In Equation 30, λ_S is the characteristic vector of S , M is the marking of net \mathcal{N} , and ξ_S is an integer such that $0 < \xi_S < M_0(S)$. However, the direct imposition of Constraint 30 on the behavior of the original net \mathcal{N} , through the theory of control-place invariants, might lead to a controlled net \mathcal{N}_c that is not PT-ordinary, and therefore, it cannot have its liveness tested according to the criterion of Theorem 3. For this reason, Constraint 30 is enforced upon the original net \mathcal{N} , through the enforcement of another inequality, of the type:

$$\kappa_S^T M_S \leq M_0(S) - \xi_S \quad (31)$$

In Equation 31, κ_S is a binary vector with its non-zero elements corresponding to some places $p \in P_S$. Hence, the inequality of Equation 31 is of the same type with that of Equation 23, and it can be enforced on the original net \mathcal{N} by the super-imposition of a monitor place such that the resulting net \mathcal{N}_c belongs to the class of ES^3PR nets. Furthermore, the support of the vector κ_S is selected in a way ensuring that (i) the satisfaction of Constraint 31 implies the satisfaction of Constraint 30, and (ii) the resulting controlled net \mathcal{N}_c does not possess any additional uncontrolled siphons; the feasibility of such a selection is established in [3] and we refer the reader to that work for the relevant details.

In addition, the result of Corollary 3, when combined with Proposition 1, indicate that it might be possible to control the entire set of siphons $S \in \mathcal{S}^{MU}$ by explicitly controlling only a subset S^E of \mathcal{S}^{MU} with $|S^E| = \text{rank}[NF(\lambda_S) : S \in \mathcal{S}^{MU}]$. The work of [8] characterizes each siphon $S \in S^E$ as “elementary”, and it proposes to satisfy the criterion of Corollary 3 for every siphon $S \in \mathcal{S}^{MU}$, by appropriately selecting (i) the set of elementary siphons $S^E \subseteq \mathcal{S}^{MU}$, and (ii) the right-hand-side vector $[\xi_S : S \in S^E]$ of Equation 30, where the latter is considered only for the elementary siphons. A last observation necessary to interpret the methodology of [8] through the result of Corollary 3, is that, in S^3PR nets, every siphon $S \in \mathcal{S}^{MU}$ has $\overline{GCM}(\lambda_S) = GCM(M_0, \lambda_S)$. Hence, [8] also proposes to ignore, in the evaluation of G'' for every siphon $S \in \mathcal{S}^{MU} \setminus S^E$, the impact of elementary siphons with negative coefficients a_i in the corresponding expansion of $NF(\lambda_S)$.

Liveness enforcing supervision of S^3PR nets based on the “mixing” of the presented approaches It is clear from the above discussion, that the methodology proposed in [8], establishes the liveness of any given S^3PR net, by enforcing an appropriately selected lower bound, ξ_S , for the compound marking of every elementary siphon $S \in S^E$. On the other hand, the algebraic LES of the type expressed by Equation 23, attain the liveness of the controlled net, by controlling the maximum compound marking of certain subsets of the place set P_S , defined by the rows of the LES-defining matrix A . One can easily envision an LES that constitutes a “mixture” of both inequality types defined by Equations 23 and 30: Under this new LES, some of the net siphons will be controlled by the LES part pertaining to Equation 23, and the rest of them will be controlled through the LES scheme established by Equation 30. Hence, it can be claimed that the results of Section III, and in particular, those of Corollaries 2 and 3,

constitute a “unifying framework” for interpreting and extending many of the past results available in the literature with respect to the liveness enforcing supervision of $(E)S^3PR$ nets. The complete characterization of the implications of this new framework, and its potential for developing ever more permissive supervisors for ES^3PR nets, while maintaining computational tractability, are important issues for further research.

V. CONCLUSIONS

This paper (i) introduced some new tests for siphon control in marked PN systems, and (ii) it demonstrated the ability of these tests to (re-)interpret, generalize and unify a number of past results regarding the deadlock avoidance and the liveness enforcing supervision of certain PN classes modelling sequential resource allocation systems. Our future work will seek to extend these results to broader PN classes, and to unravel their complete potential regarding the deadlock-freedom and liveness enforcing supervision of these nets.

REFERENCES

- [1] F. Chu and X.-L. Xie, “Deadlock analysis of petri nets using siphons and mathematical programming,” *IEEE Trans. on R&A*, vol. 13, pp. 793–804, 1997.
- [2] J. Desel and J. Esparza, *Free Choice Petri Nets*. Cambridge University Press, 1995.
- [3] J. Ezpeleta, J. M. Colom, and J. Martinez, “A petri net based deadlock prevention policy for flexible manufacturing systems,” *IEEE Trans. on R&A*, vol. 11, pp. 173–184, 1995.
- [4] M. Jeng, X. Xie, and M. Y. Peng, “Process nets with resources for manufacturing modeling and their analysis,” *IEEE Trans. on Robotics & Automation*, vol. 18, pp. 875–889, 2002.
- [5] J. Park and S. Reveliotis, “Algebraic synthesis of efficient deadlock avoidance policies for sequential resource allocation systems,” *IEEE Trans. on R&A*, vol. 16, pp. 190–195, 2000.
- [6] J. Park and S. A. Reveliotis, “Deadlock avoidance in sequential resource allocation systems with multiple resource acquisitions and flexible routings,” *IEEE Trans. on Automatic Control*, vol. 46, pp. 1572–1583, 2001.
- [7] S. A. Reveliotis, “On the siphon-based characterization of liveness in sequential resource allocation systems,” in *Applications and Theory of Petri Nets 2003*, 2003, pp. 241–255.
- [8] Z. W. Li and M. C. Zhou, “Elementary siphons of petri nets and their application to deadlock prevention in flexible manufacturing systems,” *IEEE Trans. on SMC – Part A*, vol. 34, pp. 38–51, 2004.
- [9] T. Murata, “Petri nets: Properties, analysis and applications,” *Proceedings of the IEEE*, vol. 77, pp. 541–580, 1989.
- [10] S. A. Reveliotis, *Real-time Management of Resource Allocation Systems: A Discrete Event Systems Approach*. NY, NY: Springer, 2005.
- [11] W. L. Winston, *Introduction To Mathematical Programming: Applications and Algorithms, 2nd ed.* Belmont, CA: Duxbury Press, 1995.
- [12] A. Giua, F. DiCesare, and M. Silva, “Generalized mutual exclusion constraints on nets with uncontrollable transitions,” in *Proceedings of the 1992 IEEE Intl. Conference on Systems, Man and Cybernetics*. IEEE, 1992, pp. 974–979.
- [13] J. O. Moody and P. J. Antsaklis, *Supervisory Control of Discrete Event Systems using Petri nets*. Boston, MA: Kluwer Academic Pub., 1998.
- [14] F. Tricas, F. Garcia-Valles, J. M. Colom, and J. Ezpeleta, “A structural approach to the problem of deadlock prevention in processes with resources,” in *Proceedings of the 4th Workshop on Discrete Event Systems*. IEE, 1998, pp. 273–278.