α-Siphons of a Suboptimal Control Model of a Subclass of Petri Nets

Daniel Yuh Chao
Department of Management and Information Systems, National ChengChi University, Chinese Taipei
E-mail: yuhyaw@gmail.com
Received October 12, 2010; revised January 9, 2011; accepted January 10, 2011

Abstract

It has been a hot research topic to synthesize maximally permissive controllers with fewest monitors. So far, all maximally permissive control models for a well-known benchmark are generalized Petri net, which complicates the system. In addition, they all relied on time-consuming reachability analysis. Uzam and Zhou apply First-met-bad-marking (FBM) method to the benchmark to achieve a near maximal permissive control policy with the advantage of no weighted control (WC) arcs. To improve the state of the art, it is interesting to synthesize optimal controller with as few weighted arcs as possible since it is unclear how to optimize the control for siphon involving WC arcs. This paper explores the condition to achieve optimal controller without WC and defining a new type of siphon, called α-siphon. If the condition is not met, one can apply the technique by Piroddi et al. to synthesize optimal controllers with WC.

Keywords: Petri Nets, Siphons, Controllability, FMS, S3PR

1. Introduction

Petri nets are a popular and powerful formalism to handle deadlock problems in a resource allocation system that is a technical abstraction of contemporary technical systems. Petri nets (PN) have been employed to model FMS to discover that insufficiently marked siphons cause deadlocks [1-4].

Uzam and Zhou [5] propose an iterative approach. At each iteration, a first-met bad marking (FBM) is singled out from the reachability graph of a given Petri net model. The objective is to prevent this marking from being reached via a place invariant of the Petri net. A well-established invariant-based control method is used to derive a control place. This process is carried out until the net model becomes live. The proposed method is generally applicable, easy to use, effective and straightforward although its off-line computation is of exponential complexity. Two FMS are used to show its effectiveness and applicability.

Although reaching 19 states fewer and 6 more monitors than that the optimal one by Piroddi et al. for a well-known benchmark, it does not employ weighted control arcs and runs more efficiently. Piroddi et al. [6,7] further increase it to the optimal 21581 states using the set covering approach. However, the computation is expensive since the set-covering problem involves a large system of inequalities with numerous (the number of minimal siphons) variables. Redundant monitors must be identified based on the method in [8] during each iteration, which entails exponential time complexity. Thus, the computational burden remains high and the method is not applicable to large FMS.

Furthermore, unlike that in [5], quite a few control arcs are weighted rendering the net to be a general Petri net (GPN), which are much harder to analyze than the ordinary control net by Uzam and Zhou. The traditional MIP method cannot be extended to GPN. Hence, Piroddi et al. transformed weighted arcs into ordinary ones, which sometimes may cause unnecessary deadlocks as mentioned in [5].

Our approach [9-11] categorizes SMS into basic, compound, control and mixture siphons and derives their controllability. If one carefully selects a sequence of emptiable siphons to add monitors, the number of monitors required can be reduced. Mixture siphons containing nonsharing resource places may be emptiable.

This method does not need to enumerate all minimal siphons, nor to compute the reachability graph. Also no iterations are required and no need to remove redundant
monitors. Hence, the computation burden is much less than those by Uzam et al. as well as Piroddi et al. In addition, no control arcs are weighted.

However, the resulting model of the well-known $S^{PR}$ reaches fewer (21363) states than the one (21562) in [5], but with 11 monitors and 50 control arcs fewer than 19 monitors and 120 control arcs reported in [5].

Without the knowledge of unmarked siphons, Uzam and Zhou employ a simplified generalized mutual-exclusion constraints (GMECs) equivalently setting the number of tokens in the complementary set $S$ of a siphon $S$ fewer than the initial number of tokens in $S$ by one. This excludes some live states where the number of tokens in $S$ may equal the initial number of tokens in $S$. The GMEC by Piroddi et al. sets $S$ to be always marked and does not cause states to be lost.

To avoid WC while not losing live states, we need to understand why the state loss occurs. An earlier paper does not cause states to be lost.

Section 4 develops the condition for an siphon, mixture and ordinary connected iff

$$P,T,F$$

This section focuses on estimating the number of lost states without reachability help by proposing one way to list all lost states and understand why the state loss occurs. An earlier paper

$$\text{If a marking in } R(N,M_0), \text{ a total deadlock if } \forall t \in T ; t \text{ is dead. A PN is live under } M_0 \text{ if } \forall t \in T, t \text{ is live under } M_0.$$  

For a Petri net $(N, M_0)$, a non-empty subset $S$ of places is called a siphon (trap) if $S \subseteq T (r^* \subseteq r)$, i.e., every transition having an output (input) place in $S$ has an input (output) place in $S (r)$. If $M_0(S) = \sum_{p \epsilon S} M_0(p) = 0$, $S$ is called a siphon at $M_0$. A minimal siphon does not contain a siphon as a proper subset. It is called a strict minimal siphon (SMS), if it does not contain a trap.

A P-vector (place vector) is a column vector $Y : P \rightarrow Z$ indexed by $P$ where $Z$ is the set of integers. For economy of space, we use $\sum_{p \epsilon P} L(p)p$ to denote a P-vector.

The incidence matrix of $N$ is a matrix $[N] : P \times T \rightarrow Z$ indexed by $P$ and $T$ such that $[N]^t(p, t) = F(t, p)$ and $[N](p, t) = F(p, t)$. We denote column vectors where every entry equals 0(1) by $\mathbf{0}(1)$. $Y^t$ and $[N]^t$ are the transposed versions of a vector $Y$ and a matrix $[N]$, respectively. $Y$ is a P-invariant (place invariant) if and only if $Y \neq 0$ and $Y^t [N] \neq 0$ hold. $|Y| = \{p \epsilon P | Y(p) \neq 0\}$ is the support of $Y$. A minimal P-invariant does not contain another P-invariant as a proper subset. If a siphon $S \subset \{Y\}$, then $[S] = |[Y]| \setminus S$ is called the complementary siphon of $S$ and $S \cap [S]$ is the support of a P-invariant.

Let $Y_0$ be the minimal P-invariant associated with control place $V$. $H(V) = |V| - |Y_0| \setminus \{V\}$ is called the controller (or disturbed) region of the set of all places $V$.

Definition 1 [1]: A simple sequential process (SSP) is a net $N = (P \cup \{p^0\}, T, F)$ where: 1) $P \neq \emptyset$, $p^0 \notin P$ (p^0 is called the process idle or initial or final operation place); 2) $N$ is strongly connected state machine (SM) and 3) every circuit C of $N$ contains the place $p^0$.

Definition 2 [1]: A simple sequential process with resources (SSPR), also called a working processes (WP), is a net $N = (P \cup \{p^0\}, T, F, R)$ so that 1) the subnet generated by $X = P \cup \{p^0\} \cup T$ is an SSP, 2) $R \neq \emptyset$ and $P \cup \{p_i\} \cap R = \emptyset$; 3) $\forall p \in P, \forall r \in p^0, \forall r^0 \in p^0, \exists r_p \in R, r^0 \cap P = r^0 \cap P = \{r_p\}$; 4) The two following statements are verified: $\forall r \epsilon P, a) ^+r \cap P = r^0 \cap P \neq \emptyset$; b) $^+r \cap

Copyright © 2011 SciRes.
\( r^* = \emptyset \). 5) \( \{p^0\} \cap P_r = (p^0)^* \cap P_r = \emptyset \). \( \forall p \in P, p \) is called an operation (or activity) place. \( \forall r \in P_r, r \) is called a resource place. \( H(r) = \{r \cap P \) denotes the set of holders of \( r \) (operation places that use \( r \)). \( \mathcal{H}(r) = \{r \cup H(r) = r\) and \( \mathcal{H}(r) \) is the support of a minimal P-invariant Yr that contains \( r \). Definition 3 [1]: A system of \( S^\prime \)PR (\( S^\prime \)PR) is defined recursively as follows: 1) An \( S^\prime \)PR is defined as an \( S^\prime \)PR; 2) Let \( N_i = (P_i \cup P_i^0 \cup P_{Ri} \cup T_i F_i), i \in \{1,2\} \) be two \( S^\prime \)PR so that \( (P_1 \cup P_1^0 \cap (P_2 \cup P_2^0) = \emptyset, P_{R1} \cap P_{R2} = P_C \) \( \neq \emptyset \) and \( T_1 \cap T_2 = \emptyset \). The net \( N = (P \cup P^0 \cup P_R \cup T,F) \) resulting from the composition of \( N_1 \) and \( N_2 \) via \( P_C \) (denoted by \( N_1 \circ N_2 \)), is defined as follows: 1) \( P = P_1 \cup P_2 \); 2) \( P^0 = P_1^0 \cup P_2^0 \); 3) \( P_R = P_{R1} \cup P_{R2} \); 4) \( T = T_1 \cup T_2 \) and 5) \( F = F_1 \cup F_2 \) is also an \( S^\prime \)PR. A directed circuit \( N \) is called a resource circuit, if \( \forall p \in \mathcal{G}, p \in R \). An elementary resource circuit is called a basic siphon. A compound circuit \( c \) obtained from an elementary circuit, called a compound siphon. If we add monitors to these siphons, then some new types of siphons in a certain order, then some resource or core subnets. New types (such as control siphons) of SMS can be synthesized from control subnets formed by control places. If we add monitors to these different types of siphons in a certain order, then some siphons may be redundant. We construct an SMS based on the concept of handles. Roughly speaking, a “handle” is an alternate disjoint path between two nodes. A PT-handle starts with a place and ends with a transition while a TP-handle starts with a transition and ends with a place. A core subnet can be obtained from an elementary circuit, called core circuit, by repeatedly adding handles. The control place and arcs for siphon \( S \), similar to resource places, form a number of elementary circuits. Hence, there is an elementary circuit containing adjacent control places, from which we can synthesize new problematic siphons. Definition 4: An elementary resource circuit is called a basic circuit, denoted by \( c_b \). The siphon constructed from \( c_b \) is called a basic siphon. A compound circuit \( c = c_1 \circ c_2 \circ \cdots \circ c_n \circ c_m \) is a circuit consisting of multiply interconnected elementary circuits \( c_1, c_2, \cdots, c_n \) such that \( c_1 \cap c_{i+1} \cap \{r_{pi}\}, r_{pi} \in R \) (i.e., \( c_i \) and \( c_{i+1} \) intersects at a resource place \( r \)). \( r_{pi} \) is called an inter-place. The SMS synthesized from compound circuit \( c \) (resp. control, mixture) using the Handle-Construction Procedure in [9] is called an n-compound (resp. control, mixture) siphon \( S \), denoted by \( S = S_1 \circ S_2 \circ \cdots \circ S_n \in S_m \). A siphon is called a resource siphon if it does not contain any control place. The set of compound, control, and mixture siphons for an n-compound siphon is called a family set of siphons of the n-compound siphon. Definition 5: A mixture subnet is obtained by adding non-resourceless TP-handles (containing no operation places) upon a core circuit. A siphon synthesized from a mixture subnet is called a mixture siphon. A full mixture subnet is a mixture subnet upon which we can no longer add non-resourceless TP-handles to form a larger subnet to synthesize a new siphon. Otherwise, it is called a partial mixture (briefed as p-mix) subnet. A siphon synthesized from a full (resp. partial) mixture subnet is called a full (resp. partial) mixture siphon, briefed as p-mix (resp. p-mix). \( R_S \) (resp. \( C_S \)) the set of resource (resp. control) places in \( S \). An a-siphon is a mixture one with non-sharing places.

For the benchmark in Figure 1, \( S_{11} \) is an a-siphon (where \( P_{34} \) is a non-sharing place.), whose core subnet can be obtained by adding handles \([t_1 p_{30} t_2 p_{30} t_{22} V_{11}], [t_{22} p_{32} t_1 V_{30}], [V_{20} p_{30} t_{19} p_{32} t_5 V_{16}], [V_{16} t_6 V_{11}], [p_{32} t_0 p_{43}], t_{10} V_{16}, t_{16} V_{11} \) and \( t_{42} \) to Core circuit \( c = [V_{16} t_5 V_{11} t_{20} V_{16}], c_1 = \{p_{31} t_3 p_{40} t_{30} p_{32} t_{22} t_{21} p_{31}\}, c_2 = \{p_{31} t_{20} p_{43} t_{19} p_{32} t_5 p_{34} t_{43} t_{42} t_{31}\}, c_1 \cap c_2 = \{p_{31}\}, m_0 \{p_{31}\} = 1\). Table 1 lists the controlled model by Uzam et al. based on the FBM approach.

In a mixture siphon, \( \exists t \in (S^* / S), \) *t \cap S > 1), and each firing of \( t \) may remove multiple (say \( x \)) tokens from \( S \). This is the reason that the arc from \( V_S \) to \( t \) must be weighted by \( x \) if \( M_{SD}(S) < M_{MD}(S) \). Thus, \( M_{max}(|S|) < M_{MD}(S) \). In order to avoid empty \( S \), one may set \( M_{SD}(V_S) = M_{max}(|S|) \) – 1 with ordinary control arcs.

On the other hand, for a siphon \( S \) where all non-operation places are resource ones, \( \forall t \in (S^* / S), \) *t \cap S = 1), each firing of \( t \) (called sink transitions of the siphon) removes one token from \( V_S \) and \( S \) respectively. Thus, \( M_{max}(|S|) = M_{SD}(S) \). The same holds true for a control siphon.

Based on the above discussion, there will be no live state losses if \( M_{SD}(V_S) = M_{MD}(S) \) – 1 for the resource or control siphon since state losses occur iff there are live states \( M \) such that \( M(|S|) > M_{SD}(V_S) \). Indeed, for a mixture siphon to be emitable, it must be an a-siphon.

Lemma 1: Let \( S \) be a siphon in the family set of a 2-compound siphon involved in some state loss, then \( S \) must be an a-siphon.

Proof: The state loss would not occur if no monitor is added to \( S \). The thesis holds since there is no state loss if \( S \) is not emitable and a mixture siphon is emitable and needs a monitor iff it is an a-siphon.

Monitor \( V_{17} \) is added to \( S_{11} \) to make 19 live states to be forbidden and lost via reachability analysis in \([2] \) . In the sequel, we will develop the condition for state loss for an a-siphon since other siphons in the family set of a 2-compound do not incur state loss.
Table 1. Control model by Uzam et al. for the benchmark in Figure 1.

<table>
<thead>
<tr>
<th>i</th>
<th>FBM(M_1(V_i))</th>
<th>S</th>
<th>V_r</th>
<th>V_r'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>p_{11} + 2p_{12}(2)</td>
<td>{p_1, p_3, p_{11}, p_{13}, p_{14}, p_{46}}</td>
<td>t_4</td>
<td>t_2</td>
</tr>
<tr>
<td>2</td>
<td>p_4 + 2p_{12}(2)</td>
<td>S_1</td>
<td>t_3, t_14</td>
<td>t_5, t_13</td>
</tr>
<tr>
<td>3</td>
<td>2p_1 + p_{12}(2)</td>
<td>S_2 = {p_4, p_5, p_{11}, p_{13}, p_{26}, p_{31}, p_{42}}</td>
<td>t_8, t_14</td>
<td>t_3, t_20</td>
</tr>
<tr>
<td>4</td>
<td>p_4 + 2p_{12}(2)</td>
<td>S_3 = {p_4, p_5, p_{11}, p_{13}, p_{26}, p_{31}, p_{43}}</td>
<td>t_6, t_20</td>
<td>t_6, t_19</td>
</tr>
<tr>
<td>5</td>
<td>2p_1 + p_{13}(3)</td>
<td>S_4 = {p_4, p_{12}, V_5, V_5}</td>
<td>t_6, t_20</td>
<td>t_6, t_19</td>
</tr>
<tr>
<td>6</td>
<td>2p_4 + p_{13}(2)</td>
<td>S_5 = {p_4, p_{25}, p_{35}, p_{41}}</td>
<td>t_6, t_10, t_4</td>
<td>t_6, t_14</td>
</tr>
<tr>
<td>7</td>
<td>2p_4 + p_{14} + p_{25}(4)</td>
<td>S_4 = {p_4, p_{25}, V_5, V_5}</td>
<td>t_8, t_20</td>
<td>t_3, t_18</td>
</tr>
<tr>
<td>8</td>
<td>p_4 + p_{21} + p_{32}(3)</td>
<td>S_5 = {p_4, p_{25}, p_{31}, p_{45}}</td>
<td>t_6, t_20</td>
<td>t_6, t_18</td>
</tr>
<tr>
<td>9</td>
<td>2p_4 + p_{14} + p_{45}(4)</td>
<td>S_6</td>
<td>t_6, t_10, t_20</td>
<td>t_6, t_18</td>
</tr>
<tr>
<td>10</td>
<td>p_4 + p_{14} + p_{32} + 2p_{45}(5)</td>
<td>S_7 = {p_4, p_{25}, p_{31}, p_{35}, p_{41}, p_{43}}</td>
<td>t_6, t_10, t_20</td>
<td>t_6, t_18</td>
</tr>
<tr>
<td>11</td>
<td>p_4 + p_{14} + p_{32} + 3p_{45}(5)</td>
<td>S_8 = {p_4, p_5, p_{11}, p_{13}, p_{26}, p_{31}, p_{45}, p_{46}}</td>
<td>t_6, t_10, t_20</td>
<td>t_6, t_18</td>
</tr>
<tr>
<td>12</td>
<td>p_4 + p_{14} + p_{21} + 2p_{45}(5)</td>
<td>S_9</td>
<td>t_6, t_10, t_20</td>
<td>t_6, t_18</td>
</tr>
<tr>
<td>13</td>
<td>p_4 + p_{14} + p_{21} + p_{25}(6)</td>
<td>S_{10} = {p_4, p_{25}, p_{31}, p_{35}, p_{41}, V_5}</td>
<td>t_6, t_10, t_20</td>
<td>t_6, t_18</td>
</tr>
<tr>
<td>14</td>
<td>p_4 + p_{14} + p_{21} + 2p_{35}(5)</td>
<td>S_{10}</td>
<td>t_6, t_10, t_20</td>
<td>t_6, t_18</td>
</tr>
<tr>
<td>15</td>
<td>p_4 + p_{14} + 2p_{21} + 3p_{31} + p_{35}(6)</td>
<td>S_{10}</td>
<td>t_6, t_10, t_20</td>
<td>t_6, t_18</td>
</tr>
<tr>
<td>16</td>
<td>p_4 + p_{14} + p_{12} + p_{35}(5)</td>
<td>S_{10}</td>
<td>t_6, t_10, t_20</td>
<td>t_6, t_18</td>
</tr>
<tr>
<td>17</td>
<td>p_4 + 2p_4 + p_{14} + p_{12} + p_{35} + S_{11} = {p_4, p_{11}, p_{13}, p_{14}, p_{26}, p_{31}, p_{42}, p_{45}, p_{56}, V_5, V_5, V_{16}}</td>
<td>t_6, t_10, t_20</td>
<td>t_6, t_18</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>p_4 + 2p_4 + 2p_{14} + p_{12} + 2p_{14} + p_{12} + p_{35} + p_{35}(9)</td>
<td>S_{12} = {p_4, p_{11}, p_{13}, p_{14}, p_{26}, p_{31}, p_{42}, p_{45}, p_{56}, V_5, V_{16}, V_{18}}</td>
<td>t_6, t_10, t_20</td>
<td>t_6, t_18</td>
</tr>
<tr>
<td>19</td>
<td>p_4 + 2p_4 + p_{14} + p_{12} + p_{14} + p_{12} + p_{35} + p_{35}(9)</td>
<td>S_{13} = {p_4, p_{11}, p_{13}, p_{14}, p_{26}, p_{31}, p_{42}, p_{45}, p_{56}, V_5, V_{16}, V_{17}}</td>
<td>t_6, t_10, t_20</td>
<td>t_6, t_18</td>
</tr>
</tbody>
</table>
4. Condition for State Loss

To have lost live states, some live states must be forbidden by the addition of Monitor \( V_s \). For states to be live, the \( a \)-siphon \( S \) must be always marked. For states to be forbidden, the total number of tokens in the complementary set \([S]\) of \( S \) must remain at its maximum, which cannot occur in the presence of ordinary \( V_s \). To turn \( M(S) > 0 \) (live) from \( M'(S) = 0 \) while maintaining \( M([S]) = M'(\{[S]\}) = M_{\text{max}}([S]) \) (forbidden), a token must be shifted from one place in \([S]\) to another place in \([S]\). In the sequel, we first deal with liveliness of lost states followed by two different cases where state loss may or may not occur.

Lemma 2: Let \( S \) be a siphon and \( M_d(S) > 0 \). \( M(S) > 0 \) \( \{M \in R(N, M_0)\} \), if no transitions in \( S'\setminus S \) ever fire.

Proof: Only transitions in \( S'\setminus S \) can fire to move tokens from \( S \) into \([S]\). Transitions in \( S'\setminus S \) fire to move tokens from \( S \) into \([S]\). Hence, the thesis holds.

Observation 1: Let \( S \) be an \( a \)-siphon, \( \exists V_s \in S, (S'\setminus S) \cap V_s \setminus S \neq \emptyset \).

For the \( a \)-siphon \( S = S_{23} \) in Table 1, \((S'\setminus S) \cap V_s \setminus S \neq \emptyset \). If \( t_1 \) and \( t_8 \) never fire, tokens in \( S \) cannot leak out from \( S \). There are \( 3 \) \( V_s \) in \( S_{21}, V_s, V_1 \), and \((S'\setminus S) \cap V_s \setminus S \neq \emptyset \) for \( t_1, t_8 \).

Lemma 3: Let \( S \) be an \( a \)-siphon, \( \forall V_s \in S, M(V_s) = 0 \), \( M \in R(N, M_0) \). Then no transitions in \( S'\setminus S \) can ever fire.

Proof: The thesis holds since all transitions in \( V_s \) are disabled owing to the fact that \( M(V_s) = 0 \) and \((S'\setminus S) \cap V_s \setminus S \neq \emptyset \) by Observation 1.

The above lemma help prove that markings, where an \( a \)-siphon is always marked, are live ones.

Definition 6: Let \( S \) be an \( a \)-siphon, \( RC = \{r | r \in P_b, r \in R(C) \}, C_s \in C \), the set of resource places whose holder places are also in \( \{r \in C_s \} \), \( r \) is an inter-place. \( p' = H(r_p) \) is called a skew place.

Theorem 1: Let \( S \) be an \( a \)-siphon, \( M(p) = 0 \), \( M(H(p) \cap [S]) = M(H(p)) \), \( \forall p \in [a \cup C_S] \), \( M_a \in R(N, M_0) \). Then all transitions in \( S' \) are dead.

Proof: It is easy to see that \( M_d(S) = 0 \) and all transitions in \( S' \) are dead.

For the example, \( C_S = \{V_1, V_6\} \), \( R(V_6) = \{p_{31}, p_{32}, p_{41}, p_{43}\} \), \( R_C = \{p_{30}, p_{31}, p_{32}, p_{40}, p_{41}, p_{42}, p_{43}\} \), \( p_{30} = p_{31} = p_{32} = p_{40} = p_{41} = p_{42} = p_{43} \), \( p_{30} \) and \( p_{31} \) are unmarked under FBM17. \( H(r_p) = \{p_{40}, p_{43}\} \) and \( p_{30} \) in \( [S] \) is a skew place. \( M_d(p_{32}) = M_d(H(p_{30}) \cap [S]) = M_d(p_{40}) = 1 \), \( M_d(p_{31}) = M_d(H(p_{40}) \cap [S]) = M_d(p_{43}) = 2 \), and \( M_d(p_{42}) = 1 \). The result is the same as the case in \( \{a \cup C_S\} \). Thus all output transitions of \( p \) are dead. The rest transitions are output transitions of \( p_{16}, p_{26}, p_{42}, p_{5}, p_{23}, p_{3} \), which are also dead since \( M_d(p_{16}) = M_d(p_{26}) = M_d(p_{42}) = M_d(p_{5}) = M_d(p_{23}) = 0 \).

We first add Monitor \( V' \), so that \( H(V') = \Psi \). This induces dead submarkings (markings restricted to operation places or \( \Psi \)) \( FBM_a = p_2 + p_3 + p_4 + p_5 + p_6 + p_21 + p_{22} + p_{24}, FBM_b = p_2 + p_3 + p_4 + p_5 + p_6 + p_21 + p_{22} + p_{24} \) and \( FBM_c = p_2 + p_3 + p_1 + p_6 + p_21 + p_{22} + p_{24} \), and \( M_d(V) = M_d(V_c) = 0 \). Monitors \( V_1, V_18, V_9 \) (called induced monitors) are added with \( M_d(V_1) = 9 \), \( M_d(V_18) = 9 \), and \( M_d(V_9) = 9 \), respectively.

Now Monitors \( V' \) is redundant since its controller region \( \Psi' = \{p_2, p_3, p_5, p_7, p_9, p_{21}, p_{22}, p_{24}\} \) is a subset of that \( (\Psi = \{p_2, p_3, p_4, p_5, p_7, p_9, p_{21}, p_{22}, p_{24}\}) \) for Monitor \( V_17 \) by the following lemma.

Lemma 4 [11]: Let \( S \) be an SMS. \( \delta_1 \subset \delta \subset [S] \). \( M, M_1 \) \( \in R(N, M_0) \) such that \( M(\delta) = M_1(\delta) = M_{\text{max}}([S]), V \) and \( V_1 \) and \( V_1 \) are monitors added such that \( M_d(V) = M_d(V_1) = M_{\text{max}}([S]) \) and \( [V] = \delta \), \( [V_1] = \delta_1 \). Then \( V_1 \) is redundant.

**Theorem 2**: Let \( S \) be an \( a \)-siphon, \( \Psi \) the set of marked operation places when \( S \) is unmarked under \( M_a \) and \( V_5 \) is the monitor added to \( S \) with \( M_d(V_5) = M_{\text{max}}([S]) - 1 \) and \( H(V_5) = \Psi \). Let \( M_d(p) = M_d(r) = 1 \), \( M_d(p') = M_d(p') = 1 \), \( M_d(p^*) = M_d(p^*) \), \( \forall p^* \in P \cap \{p, p'\} \), \( M_b \in R(N, M_0) \), where \( p \in H(r_p), r_p \) an inter-place, \( r \in ([\delta^*] \cap a), p \in p^* \cap H(r) \) and \( M_a \) was defined in Theorem 1. If \( H(r) \cap \{p' \} \) and \( r \in S \), then \( M_a \) is a nonlive marking. 2) There are no lost live states iff \( M_d(p) > 0 \) or \( M_d(r) = 1 \) and a monitor has been added to prevent \( M_a \) from being reached.

**Proof**: 1) Among all dead transitions under \( M_a \), only output transitions of \( r \) may be enabled under \( M_a \), since \( M_d(r) = 1 \). If \( H(r) \cap \{p' \} \) is the only possible enabled transition is the output transition of both \( r \) and \( p \). However, after \( r \) fires, it reaches \( M_a \), which is a dead marking. Thus, \( M_b \) is a nonlive marking. But \( t \) is disabled by \( V_5 \) since \( t \) is the output transition of \( V_5 \) and \( M(V_5) = 0 \). \( M_d([S]) = M_d([S]) + M_d(p) + M_d(p') - M_d(p^*) - M_d([S]) \).

2) First assume a) \( M_d(p^*) > 0 \) (or \( M_d(r) > 1 \). Let \( M_a \in R(N, M_0) \) be such that \( M_d(p^*) = M_d(p^*) + 1 \) i.e., adding a token to \( p^* \), \( M_d(p^*) = M_d(p^*) + 1 \) to ensure \( M_a(V) = 0 \), \( M_d(p^*) = M_d(p^*) \), \( \forall p^* \in P \cap \{p, p'\} \), where \( p, p^* \in H(r) \cap H(V), p' \in H(V), V \in S, p \in C_S \).

Then \( H(V) \cap S = 1 \). By Lemma 2, \( S \) remains marked since \( S \) is unmarked to disable its output transition in \( S' \setminus S \). All markings \( M' \) where \( S \) and all other siphons in the final live controlled net are marked and
$M'(p) = M(p), \forall p \in S \cup \{S\}$. Such states are live as proved below. Assume it necessarily evolves to a deadlock state $M^*$, then there exist an unmarked siphon under $M^*$, which violates the fact that all siphons have been controlled.

These states are lost since $M(t) = M(\Psi) = M_{\max}(\{V_3\})$, which are not reachable by Monitor $V_3$ with $M_0(V_3) = M_{\max}(\{S\}) - 1$.

Next consider b) $M_0(p') = 0$ (or $M_0(r) = 1$). $\Psi$ now does not include $p'$. One can no longer add a token to $p'$ to induce $M(H(r') \cap S) = 1$. Thus, there are no lost states.

a) and b) together prove the thesis.

For the example, $p' = p_{35}, r = p_{41}$, and $H(r) \cap \{S\} = \{p'\}$. $r$ has only one output transition $t_i$ with an input operation place in $\Psi$. The above theorem indicates that when $M_0(p_{41}) = 1$ (instead of 2 as in Figure 1), there will be no loss of good states. $V_3$ is not redundant and cannot be removed for the control.

a) If $M_0(p_{41}) > 1$, then there will be lost live states by adding a token to $p_{35}$ so that $M_0(p_{35}) = M_0(p_{41}) + 1$. These live states must be such that $M_0(\Psi) = M_{\max}(\Psi)$, which are not reachable by adding Monitor $V_3$ to make $M_{\max}(\Psi) - M_{\max}(\Psi)$. To make $M_0(\Psi) = M_{\max}(\Psi)$, it must be that $M_0(V_3) = M_0(V) = M(V) = 0$, $V = V_{\psi}$. This leads to $M_0(V_{\psi}) = M_0(V_{\psi}) = 5$ or

$M_0(p_4) + M_0(p_3) + M_0(p_8) = 0.$

To make $M_0(p_4) = M_0(p_3) = M_0(p_8) = 0$, implies that

$\alpha = M_0(p_4) + M_0(p_3) = M_0(p_8) = M_0(p_{22}) = 5.$

Note that the addition of Monitor $V_{\psi}$ limits $\alpha$ to be no more than 5. However, setting $\alpha$ to 5 may not make $S$ unmarked since some resource place (e.g., $p_{23}$ or $p_{35}$) in $S$ may be marked ($M(S) > 0$) even though $V_{\psi}$ is unmarked. These states $M(S) > 0$ will stay so (and are live as proved above) since transitions in $S \setminus S$ are disabled by output control arcs from unmarked $V_{\psi}$ and $V_{\psi}$. Note that $V_3$ is redundant and can be removed.

b) If $M_0(p_{41}) = 1$, then there will be no lost live states since $M_0(p_3) = 0$ and the set $\Psi$, of marked operation places under $M$, does not include $p_3$. One can no longer add a token to $\Psi$, to make $M(S) > 0$. Hence, there are no lost states.

Theorem 3: Let $S$ be an $a$-siphon, $\Psi$ the set of marked operation places when $S$ is unmarked under $M$, and $V_3$ is the monitor added to $S$ with $M_0(V_3) = M_{\max}(\Psi) - 1$ and $H(V_3) = \Psi$. Let $M_0(p) = M_0(r) = 1, M_0(p*) = M_0(p_*), \forall p* \in P \setminus\{p, p'\}, M_0(r) = 0$, where $p \in H(r), r \in \Psi$, $p' \in p^* \setminus H(r)$. If $r \in S$ and $H(r) \cap \{S\} = \{p'\}$, then 1) $M_0$ is a nonlive marking, and 2) there are no lost live states by adding a monitor to prevent $M_0$ from being reached.

Proof: 1) Similar to the proof of Theorem 2, the output transition of both $r$ and $p$ is an output transition of $V_3$ (Monitor for $S$) and disabled by unmarked $V_3$. Other output transitions $r'$ of $r$ are also disabled as explained here. If $H(r) \cap \{S\} = \{p'\}$, then $M_0(\Psi) = M_0(p') \phi (\phi = H(r) \cap \{S\} \cup \{p'\})$ is the complementary set of another siphon $S'$; the output transition set of $V_3$ (control place for $S'$) contains $r'$. $M_0(p') = 0$ implies that $M_0(\Psi) = M_0(r) = M_0(\Psi) - 1 = M_0(M_0(\Psi)) = M_0(S) - 1 = M_0(V_3)$. Thus, $V_3$ is unmarked to disable $r'$ and all possible enabled transitions are dead and $M_0$ is a nonlive marking, which needs a monitor $V'$ with $H(V')$ the set of unmarked operation places in $\{S\}$.

2) Note that $H(V')$ does not include $p'$ since $M_0(p') = 0$. Since $M_0 \phi \{p'\} + M_0(\Psi) = M_0(r)$, there is no way to add a token (to reach states forbidden by $V'$) to enable some transition. Hence, such states are nonlive and there are no lost live states.

For the example, there are two possible pairs of ($r$, $p$): 1. $(p_{35}, p_{34})$ and 2. $(p_{41}, p_{42})$ for the above theorem. For Case 1, $H(r) \cap \{S\} = \{p_{35}, p_{34}\}$. $r$ has more than one output transition ($1. t_{19} \in \{S\}$ and 2. $t_{21}$) with an input operation place in $\Psi$. $\Psi_{\psi} = H(r) \cap \{S\} \cup \{p'\} = \{p_{35}, p_{34}\}$. For Case 2, $H(r) \cap \{S\} = \{p_{41}, p_{42}\}$. $r$ has more than one output transition ($1. t_{19} \in \{S\}$ and 2. $t_{21}$) with an input operation place in $\Psi$. $\Psi_{\psi} = H(r) \cap \{S\} \cup \{p'\} = \{p_{41}, p_{42}\}$. For $\Psi_{\psi}$ the set $\Psi_b$ of marked operation places under $M_0$ does not include $p'$. One cannot add a token to a skew place in $\Psi_{\psi}$ to make $M_0(S) > 0$. Hence, there are no lost states. Combining Theorems 2 and 3, we have

Theorem 4: Let $S$ be an $a$-siphon and all necessary monitors have been added such that there are no marked set $\Psi \setminus S$ of operation places with dead output transitions. Then there are no lost live states if $M_0(p') = 0$ for all possible $p'$ (defined in Theorem 3).

Proof: Theorems 2 and 3 consider all cases where $H(r) \cap \{S\} = \{p'\}$ and $H(r) \cap \{S\} = \{p'\}$, respectively. Theorem 2 proves that there are no lost live states if $M_0(p') = 0$ for all possible $p'$. Theorem 3 proves that if $M_0(p') = 0$, then $M_0$ is a nonlive marking and there are no lost states. Similar to the proof for Theorem 2, one can show that if there are no lost states, then it must be that $M_0(p') = 0$. All cases have been considered and the thesis is proved.

In summary, this section develops the condition for a mixture siphon $S$ to be involved in reaching fewer live states. After adding a monitor $V_{\psi}$ to $S$, new unmarked siphons may be generated. One new set of unmarked operation places may cover $H(V_3)$ of $V_3$, as a proper subset. This makes $V_3$ redundant and some live states lost.

The physics of loss of live states is as follows. Adding a token to a skew place (e.g., $p_4$ in Figure 1) of $S$ reduces
a token in the holder set (e.g., \( p_{21} \) in Figure 1) of a resource place \( r \) (e.g., \( p_{32} \) in Figure 1) in \( S \), which in turn induces a token in \( r \), thus making \( S \) marked. Such a state is live and forbidden since the total number of tokens in \( \psi \) remains unchanged.

5. Conclusions

This paper enhances an earlier paper (which estimates the number of lost states without reachability analysis) and develops theory to identify the siphon responsible for lost states for a well-known benchmark and explores the condition to achieve optimal controller without WC. If the condition is not met, one can apply the technique by Piroddi et al. to synthesize optimal controllers with WC. Future work should be addressed to synthesize suboptimal controller without WC when the condition cannot be satisfied.

6. References


