Fault detection for discrete event systems using Petri nets with unobservable transitions

Alessandro Giua, Carla Seatzu

Abstract—In this paper we present an efficient approach for the fault detection of discrete event systems using Petri nets. We assume that some of the transitions of the net are unobservable, including all those transitions that model faulty behaviors. We prove that the set of all possible firing sequences corresponding to a given observation can be described as follows. First a set of basis markings corresponding to the observation are computed together with the minimal set of transitions firings that justify them. Any other marking consistent with the observation must be reachable from a basis marking by firing only unobservable transitions. For the computation of the set of basis markings we propose a simple tabular algorithm and use it to determine a basis reachability tree that can be used as a diagnoser.

I. INTRODUCTION

The diagnosis of discrete event systems is a research area that has received a lot of attention in the last years and has been motivated by the practical need of ensuring the correct and safe functioning of large complex systems. Several original theoretical approaches have been proposed [12], [6], [4], [14], [7], [9] to solve this problem.

Petri net models have often been used in this context: the intrinsically distributed nature of Petri nets where the notion of state (i.e., marking) and action (i.e., transition) is local has often been an asset to reduce the computational complexity involved in solving a diagnosis problem. Among the different contributions in this area we recall the work of Ushio et al. [13], Benveniste et al [1], [2], Jiroveanu and Boel [3], [8].

In this paper we deal with the failure diagnosis of discrete event systems modelled by place/transition nets. We assume that faults are modelled by unobservable transitions, but there may also exist other transitions that represent legal behaviors that are unobservable as well. Thus we assume that the set of transitions can be partitioned as $T = T_o \cup T_u$ where $T_o$ is the set of observable transitions, and $T_u$ is the set of unobservable transitions. The set of fault transitions is denoted $T_f$ and it holds $T_f \subseteq T_u$.

As an example consider the net in Fig. 1. The set of observable transitions is $T_o = \{t_1, t_4, t_7\}$. The set of unobservable transition is $T_u = \{t_2, t_3, t_5, t_6\}$ and, for a better understanding, an unobservable transition $t_i$ is labelled $\varepsilon_i$. The only fault transition is $t_6$. This net models a communication system: messages ready to be sent are divided into two packets (transition $t_1$) to be sent on two separate channels (place $p_4$ and $p_5$). The two packets are finally combined and an acknowledgement is sent to the sender (transition $t_7$). A fault occurs when a packet that should be travelling on the second channel is erroneously moved to the first channel (transition $t_4$). As can be seen, the fault transition $t_0$ is not observable but there exist several other unobservable transitions as well.

In this paper we extend the previous work as follows. Firstly we relax the assumption that the unobservable net be backward conflict-free. In this case the basis marking associated to a given observation $w \in T_o^*$ is not necessarily

1In Fig. 1 the unobservable subnet is acyclic because there exists no oriented cycle containing only unobservable transitions.

A net is backward conflict-free if all transitions have no output common place. In Fig. 1 the unobservable subnet is not backward conflict-free because place $p_4$ has in input two unobservable transitions, $\varepsilon_2$ and $\varepsilon_6$.\n
![Fig. 1. A net describing a communication system.](image-url)
unique any more, and we discuss how this set can be described in terms of minimal explanations\textsuperscript{3} following also the approach of Jiriveanu and Boel \cite{3,8}. A tabular algorithm for the computation of minimal explanations is also presented in the paper.

Secondly, we present an original technique to design an observer for bounded nets. We define for each observation $w$ a set $\mathcal{M}(w)$ composed of pairs $(M, y)$ where $M$ is a basis marking corresponding to $w$ and $y$, that we call its justification, is the firing vector of unobservable transitions that must have fired to reach it. We also present an algorithm for constructing a basis reachability tree (BRT); this is a deterministic automaton whose edges are labelled by the starting from any of the basis markings in $\mathcal{M}(w)$. We denote $\mathcal{M}(w)$ by the firing of either $t_4$ or $\varepsilon_2 \varepsilon_3 \varepsilon_5$.

A transition $t$ is enabled at $M$ iff $M \geq Pre(\cdot, t)$ and may fire yielding the marking $M' = M + C(\cdot, t)$. We write $M_0[\sigma]$ to denote that the sequence of transitions $\sigma = t_{j_1} \cdots t_{j_k}$ is enabled at $M_0$, and we write $M[\sigma]$ $M'$ to denote that the firing of $\sigma$ yields $M'$. Given a sequence $\sigma \in T^*$, we call $\pi : T^* \rightarrow \mathbb{N}^n$ the function that associates to $\sigma$ a vector $y \in \mathbb{N}^n$, named the firing vector of $\sigma$. In particular, $y = \pi(\sigma)$ is such that $y(t) = k$ if the transition $t$ is contained $k$ times in $\sigma$. A marking $M$ is reachable in $\langle N, M_0 \rangle$ if there exists a firing sequence $\sigma$ such that $M_0[\sigma] M$. The set of all markings reachable from $M_0$ defines the reachability set of $\langle N, M_0 \rangle$ and is denoted $R(N, M_0)$. Finally, we denote $PR(N, M_0)$ the potentially reachable set, i.e., the set of all markings $M \in \mathbb{N}^m$ for which there exists a vector $y \in \mathbb{N}^n$ that satisfies the state equation $M = M_0 + C \cdot y$, i.e., $PR(N, M_0) = \{M \in \mathbb{N}^m \mid \exists y \in \mathbb{N}^n : M = M_0 + C \cdot y\}$. It holds that $R(N, M_0) \subseteq PR(N, M_0)$.

A Petri net having no directed circuits is called acyclic. For this subclass the following result holds.

**Theorem 2.1:** \cite{5} Let $N$ be an acyclic Petri net.

(i) If the vector $y \in \mathbb{N}^n$ satisfies the equation $M_0 + C \cdot y \geq 0$ there exists a firing sequence $\sigma$ firable from $M_0$ such that the firing vector associated to $\sigma$ is equal to $y$.

(ii) A marking $M$ is reachable from $M_0$ iff there exists a non negative integer solution $y$ satisfying the state equation $M = M_0 + C \cdot y$, i.e., $R(N, M_0) = PR(N, M_0)$.

A net system $\langle N, M_0 \rangle$ is bounded if there exists a positive constant $k$ such that, for $M \in R(N, M_0)$, $M(p) \leq k$. A net is said structurally bounded it is bounded for any initial marking.

A labeling function $L : T \rightarrow E \cup \{\varepsilon\}$ assigns to each transition $t \in T$ either a symbol from a given alphabet $E$ or the empty string $\varepsilon$.

We denote as $T_u$ the set of transitions whose label is $\varepsilon$, i.e., $T_u = \{t \in T \mid L(t) = \varepsilon\}$. Transitions in $T_u$ are called unobservable or silent.

In this paper we assume that the same label $\varepsilon \in E$ cannot be associated to more than one transition. Thus, being the labeling function restricted to $T_u = T \setminus T_o$, an isomorphism, with no loss of generality we assume $E = T_o$. Transitions in $T_o$ are called observable.

In the following we denote as $C_u (C_o)$ the restriction of the incidence matrix to $T_o (T_u)$.

We denote as $w$ the word of events associated to the sequence $\sigma$, i.e., $w = L(\sigma)$. Note that the length of a sequence $\sigma$ (denoted $|\sigma|$) is always greater or equal than the length of the corresponding word $w$ (denoted $|w|$). In fact, if $\sigma$ contains $k'$ transitions labeled $\varepsilon$ then $|\sigma| = k' + |w|$.

Moreover, we denote as $s_{\sigma_o}$ the sequence of null length and $\varepsilon$ the empty word. We use the notation $w_i \preceq w$ to denote the generic prefix of $w$ of length $i \leq k$, where $k$ is the length of $w$.

**Definition 2.2:** Given a net $N = (P, T, Pre, Post)$, and a subset $T' \subseteq T$ of its transitions, we define the $T'$—induced subnet of $N$ as the new net $N' = (P, T', Pre', Post')$.
where $Pre', Post'$ are the restriction of $Pre, Post$ to $T'$. The net $N'$ can be thought as obtained from $N$ removing all transitions in $T \setminus T'$. We also write $N' \prec_T N$. □

III. MINIMAL EXPLANATIONS

In this section we provide some basic definitions that will be useful in the following.

**Definition 3.1:** Given a marking $M$ and an observable transition $t \in T_o$, we define

$$\Sigma(M, t) = \{\sigma \in T_o^* | M(\sigma)M', M' \geq Pre(\cdot, t)\}$$

the set of explanations of $t$ at $M$, and we denote

$$Y(M, t) = \{y \in \mathbb{N}^n | \exists \sigma \in \Sigma(M, t) : \pi(\sigma) = y\}$$

the corresponding set of firing vectors.

Thus $\Sigma(M, t)$ is the set of unobservable sequences whose firing at $M$ is necessary to enable $t$.

Among the above sequences we want to select those whose firing vector is minimal, that we call minimal explanations.

**Definition 3.2:** Given a marking $M$ and a transition $t \in T_o$, we define

$$\Sigma_{min}(M, t) = \{\sigma \in \Sigma(M, t) | y = \pi(\sigma), \not\exists \sigma' \in \Sigma(M, t) : \pi(\sigma') \preceq y\}$$

the set of minimal explanations of $t$ at $M$, and we denote

$$Y_{min}(M, t) = \{y \in \mathbb{N}^n | \exists \sigma \in \Sigma_{min}(M, t) : \pi(\sigma) = y\}$$

the corresponding set of firing vectors.

Similar definitions have also been given in [3], [8].

**Example 3.3:** Let us consider again the net in Fig. 1. Let $M_0$ be the marking shown in figure. Then $\Sigma(M_0, t_1) = \{\varepsilon\}$, namely the empty word, and $Y_{min}(M_0, t_1) = \{\emptyset\}$. In fact, $t_1$ is enabled at $M_0$ and no unobservable transition is necessary to fire $t_1$.

If we consider transition $t_7$, then $\Sigma(M_0, t_7) = \emptyset$, thus also $Y_{min}(M_0, t_7) = \emptyset$. In fact, $t_7$ is not enabled at $M_0$, and no sequence of unobservable transitions may enable it.

Now, let $M_1 = \{0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0\}$. Then $\Sigma(M_1, t_4) = \Sigma_{min}(M_1, t_4) = \{\varepsilon_2\}$. Then, let $M_2 = \{0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0\}$. Then $\Sigma(M_2, t_4) = \Sigma_{min}(M_2, t_4) = \{\varepsilon_2, \varepsilon_3 \varepsilon_6\}$. Finally, let $M_3 = \{0 \ 1 \ 1 \ 0 \ 1 \ 0 \ 0\}$. Then $\Sigma(M_3, t_4) = \Sigma_{min}(M_3, t_4) = \{\varepsilon_2, \varepsilon_6, \varepsilon_3 \varepsilon_6\}$, while $\Sigma_{min}(M_3, t_4) = \{\varepsilon_2, \varepsilon_6\}$. □

In [5] we proved the following important result.

**Theorem 3.4:** [5] Let $N = (P, T, Pre, Post)$ be a Petri net with $T = T_o \cup T_u$. If the $T_u$-induced subnet is acyclic and backward conflict-free, then $Y_{min}(M, t) = 1$.

Different approaches can be used to compute $Y_{min}(t, M)$, e.g., see [3], [8].

In this paper we suggest an approach that terminates finding all vectors in $Y_{min}(M, t)$ if applied to nets whose $T_u$-induced subnet is acyclic. It simply requires algebraic manipulations, and is inspired by the procedure proposed by Martinez and Silva [10] for the computation of minimal P-invariants. It can be briefly summarized by the following algorithm.

**Algorithm 3.5:** [Computation of $Y_{min}(M, t)$]

1. Let $\Gamma := \frac{C_u^T}{A^T}I_{n_u \times n_u}B$
   where $A^T := M - Pre(\cdot, t)$, $B := 0_{n_u}$.
2. If $A \geq 0$, goto 8, else goto 3.
3. Choose an element $A(i^*, j^*) < 0$.
4. Let $I^+ = \{i | C_u^T(i, j^*) > 0\}$.
5. If $I^+ = \emptyset$, remove the row $[A(i^*, \cdot) \mid B(i^*, \cdot)]$ from the table and goto 2.
6. For all $i \in I^+$, add to $[A \mid B]$ a new row $[A(i^*, \cdot) + kC_u^T(i, \cdot) \mid B(i^*, \cdot) + k\tilde{e}_i^T]$ where $\tilde{e}_i$ is the $i$-th canonical basis vector and $k$ is the minimum integer such that $A(i^*, j^*) + kC_u^T(i, j^*) \geq \tilde{0}^T$.
7. Remove the row $[A(i^*, \cdot) \mid B(i^*, \cdot)]$ from the table and goto 2.
8. Remove from $B$ any row that covers other rows.
9. Each row of $B$ is a vector in $Y_{min}(M, t)$.

**Example 3.6:** Let us consider again the net in Fig. 1. Let $M = [0 \ 1 \ 1 \ 0 \ 1 \ 0 \ 0]^T$ and $t = t_4$. Being

$$C_u = \begin{bmatrix} \varepsilon_2 & \varepsilon_3 & \varepsilon_5 & \varepsilon_6 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad Pre(\cdot, t_4) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

we first assume

$$\Gamma := \begin{bmatrix} 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

thus there is only one element of $A$, namely $A(1, 4)$, that is negative. Moreover, $I^+ = \{1, 4\}$. Using Algorithm 3.5 we add the following two new rows to $\Gamma$:

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and obtained from the first row of $A$ by adding the first and the fourth row of $\Gamma$, respectively. Finally, we remove the row $\Gamma(5, \cdot)$ from the table and we stop because all elements of $A$ are non negative.

Because no line covers the other, we conclude that both rows of $B$, namely

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
are elements of $Y_{\min}(M, t)$.

This result is in accordance with the previous Example 3.3, being $\Sigma_{\min}(M, t) = \{\varepsilon_2, \varepsilon_6\}$. ■

IV. BASIS MARKING

In [5] we introduced the notion of basis marking.

**Definition 4.1:** [5] Let $\langle N, M_0 \rangle$ be a net system whose unobservable subnet is backward conflict-free. Given an observation $w$, the basis marking $M_{0, w}$ is the marking reached from $M_0$ by firing $w$ and all those unobservable transitions that are strictly necessary to enable $w$. ■

The backward conflict-free assumption ensures the uniqueness of $M_{0, w}$, for any initial marking $M_0$ and any observation $w$ [5].

If the backward conflict-free assumption is relaxed, the basis marking may be not unique. This trivially follows from the simple observation that, given a marking $M$ and an observable transition $t$, the set of minimal explanations of $t$ at $M$ is generally not a singleton.

Now, in order to generalize the notion of basis marking, we introduce the following recursive definition.

**Definition 4.2:** Let $\langle N, M_0 \rangle$ be a net system where $N = \langle P, T, Pre, Post \rangle$ and $T = T_o \cup T_n$.

Let $\mathcal{M}(w) = \{(M_0, 0)\}$ and $\forall w \in T_o^*, \forall t \in T_o$, let

$$\mathcal{M}(w) = \{(M', y) \in \mathbb{N}^m \times \mathbb{N}^{p_u} |$$
$$\exists (M', y') \in \mathcal{M}(w),$$
$$y' \in Y_{\min}(M', t), y = y' + y'' \in M_0 + C(\cdot, t) + C_u y\}.$$

Finally, $\forall w \in T_o^*$, let $\mathcal{M}(w) \subseteq \mathcal{M}(w)$ such that

$$\mathcal{M}(w) = \{(M, y) \in \mathcal{M}(w) |$$
$$\exists (M', y') \in \mathcal{M}(w) : y' \leq y\}.$$

All markings $M$ such that $(M, y) \in \mathcal{M}(w)$ are called *basis marking* and the vectors $y$ are the corresponding justifications. ■

Therefore, for any observation $w$, $(M, y) \in \mathcal{M}(w)$ is a couple (marking, firing vector) such that $M$ can be reached from $M_0$ firing a sequence $\sigma$ such that $L(\sigma) = w$ and $\pi(\sigma) = \pi(\cdot) + y$. Clearly, when no observation has occurred (i.e., $w = \varepsilon$), $\mathcal{M}(w)$ is a singleton and $M = M_0$, $y = \emptyset$.

Note that each set $\mathcal{M}(w)$ only contains couples $(M, y)$ whose justifications are minimal because $\mathcal{M}(w)$ is obtained by $\mathcal{M}(w)$ removing all couples whose justifications are not minimal.

**Example 4.3:** Let us consider the net in Fig. 1. Assume that the initial marking is that shown in figure.

Let $w = t_1$. Being $Y_{\min}(t_1, M_0) = \{0\}$, if we denote as

$$M_1 = M_0 + C_o \pi(t_1) = [0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 ]^T,$$

then $\mathcal{M}(t_1) = \{M_1, 0\}$, and the null vector is the only justification of $w = t_1$ at the initial marking.

Now, assume that $t_4$ is observed, thus $w = t_1 t_4$. In such a case $Y_{\min}(M_1, t_4) = \{y_1, y_2\}$ where $y_1 = \pi(\varepsilon_2)$ and $y_2 = \pi(\varepsilon_3 \varepsilon_6)$. Now, if we denote

$$M_2 = M_1 + C_o \pi(t_4) + C_u y_1 = M_0 + C_o \pi(w) + C_u y_1 = [0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 ]^T,$$

$$M_3 = M_1 + C_o \pi(t_4) + C_u y_2 = M_0 + C_o \pi(w) + C_u y_2 = [0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 ]^T,$$

then

$$\mathcal{M}(t_1 t_4) = \{M_2, y_1, (M_3, y_2)\}.$$

Finally, assume that $t_7$ fires, thus $w = t_1 t_4 t_7$. It holds that $Y_{\min}(M_2, t_7) = \{\pi(\varepsilon_3 \varepsilon_5)\}$ and $Y_{\min}(M_3, t_7) = \emptyset$. In fact, the firing of $\varepsilon_3 \varepsilon_5$ enables $t_7$ at $M_2$, while $t_7$ is not enabled at $M_3$ and no sequence of unobservable transitions may enable it. Therefore,

$$\mathcal{M}(t_1 t_4 t_7) = \{M_4, y_1 + y_3\},$$

where

$$M_4 = M_2 + C_o \pi(t_7) + C_u y_3 = M_0 + C_o \pi(w) + C_u (y_1 + y_3) = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 ]^T = M_0.$$

The following theorem proves that our approach based on basis markings is able to characterize completely the reachability set under partial observation.

**Theorem 4.4:** Let us consider a net system $\langle N, M_0 \rangle$ whose unobservable subnet is acyclic. The following two assertions are equivalent.

1) There exists $\hat{\sigma} \in T^*$ such that $M_0[\hat{\sigma}] \hat{M}$ with $L(\hat{\sigma}) = w$ and $\pi(\hat{\sigma}) = \hat{y}$.

2) There exists $(M, y) \in \mathcal{M}(w)$ and $\sigma'' \in T_u^*$ such that $M[\sigma''] \hat{M}$ with $\hat{y} = \pi(w) + y + \pi(\sigma'')$.

**Proof:** We prove this result by induction on the length of the observed string $w$.

**(Basis step)** For $w = \varepsilon$ the results obviously holds.

**(Inductive step)** Assume the result holds for $w$. We prove it holds for $w = vt$.

Firstly, we prove 1) $\Rightarrow$ 2). In fact, if 1) holds then there exist sequences $\sigma'$ and $\sigma''$ such that

$$M_0[\sigma'] M'[t] M''[\sigma''] \hat{M}$$

where $L(\sigma') = v$, and $\sigma'' \in T_u^*$. By induction, there exists

$$(M, y) \in \mathcal{M}(v)$$

such that

$$M_0[\sigma''] M'[t] M''[\sigma''] \hat{M}$$

where $L(\sigma'') = v$, $\pi(\sigma'') = \pi(v) + y$ and $\sigma'' \in T_u^*$. Now there exists a minimal explanation $\sigma'_v \in \Sigma(M, t)$ such that $\pi(\sigma'_v) \leq \pi(\cdot)$, and being the $T_u$-induced subnet acyclic,

$$M_0[\sigma''] M[\sigma'_v] M'[t] M''[\sigma''] \hat{M}$$

where $\pi(\sigma'_v) + \pi(\sigma'') = \pi(\cdot)$ and $(M'_v, \pi(\sigma'_v)) \in \mathcal{M}(vt) = \mathcal{M}(w)$. This proves the result.
Secondly, we prove $2) \Rightarrow 1)$. In fact if $2)$ holds then there exists $\sigma' \in T^*$ such that $M_0[\sigma']M[\sigma'']M$ with $L(\sigma') = vt = w$ and hence $M_0[\sigma']\tilde{M}$ with $\sigma = \sigma'\sigma''$.

Note that this implication still holds even if the unobservable subnet is not acyclic. □

V. OBSERVER DESIGN BASED ON THE BASIS REACHABILITY TREE

In this section we focus our attention on bounded Petri nets and propose an original technique to design an observer to be used in the context of failure diagnosis.

The proposed approach consists in the design of a deterministic graph, that we call basis reachability tree (BRT).

Let us first introduce the following definitions. Let

$\mathcal{M}_0(w) = \{M \in \mathbb{N}^n \mid \exists y \in \mathbb{N}^{n_u} : (M, y) \in \mathcal{M}(w)\}$

be the set of basis markings at $w$. Let

$O(N, M_0) = \{w \in T_\sigma^* \mid \exists \sigma \in T^*, \ M_0[\sigma], \ L(\sigma) = w\}$

be the set of observable words of $(N, M_0)$.

We denote

$O_{\text{min}}(N, M_0) = \{w \in O(N, M_0) \mid \exists \tilde{w}' \in O(N, M_0) : w' < w, \ \mathcal{M}_b(w) = \mathcal{M}_b(w')\}$

the set of observable words of minimal length to which it correspond a different set of basis markings.

The BRT has as many nodes as the cardinality of $O_{\text{min}}(N, M_0)$. Each node coincides with a different set $\mathcal{M}(w)$ and each arc is labeled with an observable transition. More precisely, the BRT is an automaton on the alphabet $T_\sigma$ whose initial state is $\mathcal{M}_0 = \mathcal{M}(\varepsilon)$, and if $\delta$ is its transition function, it holds $\delta(\mathcal{M}_0, w) = \mathcal{M}(w)$ for any word $w \in O_{\text{min}}(N, M_0)$. In other words, if $w \in O_{\text{min}}(N, M_0)$, then there exists an oriented path labeled $w$ from the root node $\mathcal{M}_0$ to the node $\mathcal{M}(w)$.

The BRT of a bounded net system $(N, M_0)$ can be constructed using the following algorithm where we denote as $\mathcal{M}_b$, $\mathcal{M}_{b'}, \mathcal{M}_b, \mathcal{M}_b'$ the set of basis markings relative to the set $\mathcal{M}$ (resp., $\mathcal{M}', \mathcal{M}$.)

Algorithm 5.1: [Basis reachability tree]

1. Label the initial node $\mathcal{M}_0 = \mathcal{M}(\varepsilon)$ as the root and assign no tag to it.

2. If nodes with no tag exist, select a node $\mathcal{M}$ with no tag and:

2.1 if $\forall M \in \mathcal{M}_b$ and $\forall t \in T_\sigma$, $Y_{\text{min}}(M, t) = \emptyset$, tag $\mathcal{M}$ “dead” and go to step 2.

2.2 if $\forall t \in T_\sigma : \{M \in \mathcal{M}_b \mid Y_{\text{min}}(M, t) \neq \emptyset\} \neq \emptyset$

2.2.1 let $\mathcal{M} = \emptyset$

2.2.2 for all $(M, y) \in \mathcal{M}$

2.2.2.1 for all $\tilde{y} \in \mathcal{M}(M, t)$

2.2.2.2 compute $M' = M + C_o\pi(t) + C_u\tilde{y}$

2.2.2.3 let $\tilde{\mathcal{M}} = \mathcal{M} \cup \{M', y\}$

2.3 let $\mathcal{M}' = \{(M, y) \in \mathcal{M} \mid \tilde{y}(M', y') \in \tilde{\mathcal{M}} : y' \leq y\}$

2.4 add a new node $\mathcal{M}'$ to the graph and an arc $t$ from $\mathcal{M}$ to $\mathcal{M}'$

2.5 if already $\exists$ a node $\mathcal{M}'$ in the graph such that $\mathcal{M}_b = \mathcal{M}_b'$, tag the new node “dup”. ■

Example 5.2: The BRT of the net in Fig. 1 is reported in Fig. 2. By looking at this graph we find out all the results already discussed in the Example 4.3.

One final remark about the BRT. In the standard construction of a PN reachability/coverability graph, after a tree has been constructed, by merging identical nodes one obtains a graph that may also contain cycles. In the case of the BRT the construction of a graph is not meaningful because two nodes may correspond to the same set of basis marking but have different justifications.

Consider as an example, the net in Fig. 1 and its BRT in Fig. 2. The words $\varepsilon, t_1t_4t_7, (t_1t_4t_7)^2, \ldots$, all correspond to the same basic marking $M_0 = [1 0 0 0 0 0]^T$ but they have different justifications $0, y_1 + y_3, 2y_1 + 2y_3, \ldots$. In fact, each time the cycle $M_0[t_1t_4t_7]M_0$ the justification increases of the quantity $y_1 + y_3$.

Thus we keep the tree as it is, but to compute the set $\mathcal{M}(w)$ for a word $w$ of arbitrary length we need to keep in mind that whenever a leaf is reached, we need to continue the production from the ancestor node corresponding to the same set of basis marking while adding, each time the cycle is repeated, the corresponding justification.

VI. DIAGNOSIS

The formalism described in the previous sections for marking estimation can be used to design a diagnoser. Let us first define

$L(w) = \{\sigma \in T^* \mid M_0[\sigma], L(\sigma) = w\}$

the set of firing sequences consistent with $w \in T_\sigma^*$.

Definition 6.1: A diagnoser is a function $\Delta : T_\sigma^* \times T_f \rightarrow \{0, 1, 2, 3\}$ that associates to each observation $w$ and to each fault transition $t_f$ a diagnosis state.

$\Delta(w, t_f) = 0$ if for all $\sigma \in L(w)$ it holds that $t_f \notin \sigma$. In such a case the fault cannot have occurred because there exist no firable sequence containing $t_f$ and consistent with the observation.
Thus we have the following result.

Proposition 6.2: Consider an observed word $w \in T_0^*$. If $\Delta(w, \tau_f) = 1$ then there exists a $\sigma \in \mathcal{L}(w)$ such that $\tau_f \in \sigma$ but for all pairs $(M, y) \in \mathcal{M}(w)$ it holds that the justification $y$ of the basis marking $M$ is such that $y(\tau_f) = 0$. In such a case the fault may have occurred but not while reaching a basis marking.

$\Delta(w, \tau_f) = 2$ if there exists a pair $(M, y) \in \mathcal{M}(w)$ such that $y(\tau_f) > 0$. In such a case the fault may have occurred while reaching a basis marking.

$\Delta(w, \tau_f) = 3$ if for all $\sigma \in \mathcal{L}(w)$ it holds that $\tau_f \in \sigma$. In such a case the fault must have occurred because all firable sequence consistent with the observation contain $\tau_f$.

The diagnosis states 1 and 2 correspond both to cases in which the fault may have occurred but has not necessarily occurred. The main reason to distinguish between them is the following. In the state 1 the observed behavior does not suggest a fault has occurred, while in the state 2 at least one of the justifications for the observed behavior implies that the fault has occurred.

The diagnosis state associated to an observation $w$ can be easily computed using the BRT. We present a series of results whose proofs are rather elementary and are not given here for sake of brevity.

Let us recall that the BRT is an automaton on the alphabet $T_o$. The initial state is $M_0 = \{(M_0, \emptyset)\}$, and if $\delta$ is its transition function, it holds $\delta(M_0, w) = \mathcal{M}(w)$.

Proposition 6.3: Consider an observed word $w \in T_0^*$ such that for all $(M, y) \in \mathcal{M}(w)$ it holds $y(\tau_f) = 0$. If $\Delta(w, \tau_f) = 0$ then there does not exist a sequence $\sigma \in T_o^*$ such that $M[\sigma]$ and $\tau_f \in \sigma$. If $\Delta(w, \tau_f) = 1$ then at least one $(M, y) \in \mathcal{M}(w)$ and a sequence $\sigma \in T_o^*$ such that $M[\sigma]$ and $\tau_f \in \sigma$.

If the uncontrollable subnet is acyclic the reachability of the uncontrollable subnet can be characterized by the state equation and there exists a sequence containing transition $\tau_f$ firable from $M$ on the uncontrollable subnet if and only if the following integer constraint set (ICS) admits a solution:

$$M + C_u z \geq \bar{0}, \quad z(t_f) > 0, \quad z \in \mathbb{N}^{n_u}.$$  \hspace{1cm} (1)

Thus we have the following result.

Proposition 6.4: For a Petri net whose uncontrollable subnet is acyclic, let $w \in T_o^*$ be an observed word such that for all $(M, y) \in \mathcal{M}(w)$ it holds $y(\tau_f) = 0$. If $\Delta(w, \tau_f) = 0$ if $\forall (M, y) \in \mathcal{M}(w)$ the ICS (1) does not admit a solution:

$$\Delta(w, \tau_f) = 1$$ if $\exists a (M, y) \in \mathcal{M}(w)$ such that (1) admits a solution.

VII. CONCLUSIONS

In this paper we dealt with the problem of fault detection for discrete event systems. An original approach is presented using Petri nets with unobservable transitions. In particular, faults are modeled as unobservable transitions, and legal behaviors as well may be modeled as unobservable transitions. We first provide a characterization of the firing sequences corresponding to a given observation based on the notion of basis markings and justifications. For the computation of the set of basis markings we propose a simple tabular algorithm and use it to determine a deterministic automaton, that we call basis reachability tree, that can be used as a diagnoser.

REFERENCES


