Technical report 06-020

The on-line diagnosis of time Petri nets based on partial orders*

G. Jiroveanu, B. De Schutter, and R.K. Boel

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
phone: +31-15-278.24.73 (secretary)
URL: https://www.dcsc.tudelft.nl

* This report can also be downloaded via https://pub.bartdeschutter.org/abs/06_020.html

The on-line diagnosis of Time Petri Nets based on partial orders

G. Jiroveanu* — **B. De Schutter**** — **R.K. Boel***

* *EESA - SYSTeMS Research Group*
Ghent University
Technologiepark 914
Zwijnaarde 9052, Belgium

{george.jiroveanu,rene.boel}@ugent.be

** *Delft Center for Systems and Control (DCSC)*
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands

{b.deschutter}@tudelft.nl

ABSTRACT. *In this paper we propose an on-line diagnosis algorithm for Time Petri Nets (TPN). The plant observation is given by a subset of transitions and the faults are modeled by a subset of unobservable transitions. The plant behavior is derived on-line and the diagnosis is obtained checking whether or not some or all of the traces in the behavior that obey the plant observation contain fault events. We calculate the legal plant behavior as a set of configurations in the net unfolding. We calculate the set of legal traces in the TPN deriving for each configuration the solution set of a system of $(max, +)$-linear inequalities called the characteristic system of the configuration. We present two methods to derive the entire set of solutions of a characteristic system: the first method is based on Extended Linear Complementarity Problem while the second method is based on constraint propagation.*

RÉSUMÉ.

KEYWORDS: *discrete events systems, partial orders, on-line monitoring*

MOTS-CLÉS :

1. Introduction

This paper provides a survey of recent work on the on-line diagnosis of Time Petri Nets (TPNs) based on partial orders. In [JIR 06b] and [JIR 06c] we have described algorithms for solving this problem. Further details can be found in [JIR 06a].

TPNs are extensions of untimed Petri Nets (PNs) where timing information about the execution of some operations in the plant is available. In a TPN a transition can be fired after a delay within a given interval and its execution takes no time to complete [MER 74]. A trace in the plant comprises the transitions (events) that are executed in the TPN model (the untimed support) as well as the time of their occurrence.

In this paper we consider the plant observation given by a subset of transitions whose occurrence is always reported including also the accurate time when an observed transition is executed, measured according to a global clock. The unobservable events are silent, i.e. the execution of an unobservable transition is not acknowledged to the monitoring system. The fault transitions are modeled by a subset of the unobservable transitions.

The model-based diagnosis for TPNs requires to detect the occurrence of a fault event based on the model and the observation generated by the plant up to the current time. The on-line diagnosis requires first to calculate the set of traces that are legal, according to the model specifications, starting from the initial marking, and that obey the received observation. Then the diagnoser must check whether some or all of the legal traces include fault transitions.

There are several approaches to this problem. A possible approach is to derive off-line the full behavior of the plant and then to take into account the received observation by eliminating traces that do not obey the observation. This is a very expensive method since calculations are performed first and then discarded and more importantly it can be applied only to models of small size that moreover do not change often their structure.

In this paper we consider a different approach. When the process starts we derive time interval configurations in the TPN model up to the first discarding time. A discarding time is the time when in absence of any observation one can discard untimed support traces and it corresponds with the smallest value of the latest time when an observable event is forced to happen. The occurrence of an observable transition before the first discarding time is taken into account by eliminating traces that are not consistent with the received observation. Then the plant behavior is derived up to a next discarding time.

This method obviously requires less calculations but it requires the assumption that the faults are not predictable. This simply means that one cannot predict for sure that a fault will happen in the future. This assumption is trivial for untimed PNs. However for TPN models this assumption cannot be checked unless the full state space of the TPN under investigation is generated. Notice that if the faults can be predictable at a certain time, given the observation generated by the plant up to that time, then the on-line diagnosis method that we propose would not predict the sure occurrence of a fault at the earliest time possible.

In this paper we impose a structural condition that assures that the faults are unpredictable, namely for each fault transition there is a non-fault transition that has its pre-set included in or equal to the pre-set of the fault transition and moreover has a lower bound of its static interval that is not greater than the upper bound of the static interval of the fault transition. Notice that this condition is only a sufficient condition for the faults to be unpredictable.

The analysis of Petri Nets (PN) is an NP-hard problem because of the state space explosion due to the interleaving of concurrent events. The same problem remains also for PN models where the time is considered as a quantifiable and continuous variable. To cope with this difficulty methods based on partial orders were proposed for the analysis of untimed PNs [MCM 92],[ESP 94], [BEN 03] as well as for Time Petri Nets [SEM 96],[AUR 97],[CHA 05].

The plant analysis is based on time configurations (time-processes in [AUR 97]). A time configuration is an untimed configuration (a configuration in the net-unfolding of the untimed PN support of the TPN model) with a valuation of the execution times for its events. A time configuration is legal if there is a time trace in the original TPN that can be obtained from a linearization of the events of the configuration where the occurrence times of the transitions in the trace are identical with the valuation of their images in the time configuration. A linearization of the events in a configuration is a trace that comprises all the events of the configuration executed once s.t. the partial order between the events in the configuration is preserved in the order in which they appear in the trace.

The set of all legal time-traces in the original TPN can be obtained by computing for each configuration the entire solution set of a $(max, +)$-system of linear inequalities called the characteristic system of the configuration. The characteristic system of a configuration comprises $(max, +)$-inequalities relating the execution times of the events within the configuration as well as $(max, +)$-inequalities that assure that a conflicting event (an event that is not considered in the configuration but has its preset of conditions in the set of conditions of the configuration) was not forced to be executed.

The calculations involve time interval configurations. A time interval configuration is an untimed configuration endowed with time intervals for the execution of the events within the configuration. A time interval configuration is legal if for every event and for every execution time of the event within its execution time interval there exists a legal time configuration that considers the event executed at that time.

Thus, we need to derive for each configuration the entire solution set of its characteristic system. The naive approach to enumerate all the possible $\max$-elements would imply to interleave concurrent events which is exactly what we wanted to avoid by using partial orders to represent the plant behavior. To cope with this difficulty we present two methods that avoid the explicit consideration of all the cases for each $\max$-term in the characteristic system.

The first method uses the Extended Linear Complementarity Problem (ELCP) [DES 95] for deriving the set of all solutions of the characteristic system of the configuration. The solution set can be represented as a union of faces of a polyhedron that satisfy a cross-complementarity condition.

The second method is based on constraint propagation and exploits the partial order relation between the events within the configuration. We derive for each untimed configuration a set of hyperboxes of dimension equal to the number of events within the configuration such that the union of all the subsets of solutions that are circumscribed by the hyperboxes is a cover of the solution set.

The paper is organized as follows. In Section 2 we provide definitions and the notation used in the paper and in Section 3 we formalize the diagnosis problem for TPNs models. The analysis of TPNs based on partial orders is described in Section 4. Section 5 and Section 6 present the two methods to derive the solution set of a characteristic system of a configuration and in Section 7 we present the on-line diagnosis algorithm that we propose. The paper is concluded in Section 8 with final remarks and future work.

2. Notation and definitions

2.1. *Petri nets*

A Petri Net is a structure $\mathcal{N} = (\mathcal{P}, \mathcal{T}, F)$ where $\mathcal{P}$ denotes the set of $\mid \mathcal{P} \mid$ places, $\mathcal{T}$ denotes the set of $\mid \mathcal{T} \mid$ transitions, and $F = Pre \cup Post$ is the incidence function where $Pre(p, t) : \mathcal{P} \times \mathcal{T} \rightarrow \{0, 1\}$ and $Post(t, p) : \mathcal{T} \times \mathcal{P} \rightarrow \{0, 1\}$ are the *pre-* and *post-incidence function* that specify the arcs.

We use the standard notations: $p^\bullet$, ${}^\bullet p$ for the set of input, respectively output transitions of a place; similarly ${}^\bullet t$ and $t^\bullet$ denote the set of input places to t, and the set of output places of t respectively. A *marking* M of a PN is represented by a $\mid \mathcal{P} \mid$-vector, $M : \mathcal{P} \rightarrow \mathbb{N}$, that assigns to each place of $\mathcal{N}$ a non-negative number of tokens.

The set $\mathcal{L}_{\mathcal{N}}(M_0)$ of all legal traces of a PN, $\langle \mathcal{N}, M_0 \rangle$, with initial marking M_0 is defined as follows. A transition t is *enabled* at the marking M if $M \geq Pre(\cdot, t)$. Firing, an enabled transition t consumes $Pre(p, t)$ tokens in the input places $p \in {}^\bullet t$ and produces $Post(t, p)$ tokens in the output places $p \in t^\bullet$. The next marking is $M^{'} = M + Post(t, \cdot) - Pre(\cdot, t)$. A trace τ is defined as $\tau = M_0 \xrightarrow{t_1} M_1 \xrightarrow{t_2} \ldots \xrightarrow{t_k} M_k$, where for $i = 1 \ldots k$, $M_{i-1} \geq Pre(t_i)$. $M_0 \xrightarrow{\tau} M_k$ denotes that the sequence τ may fire at M_0 yielding M_k. Given a marking M denote by $Enbl(M)$ the set of transitions that are enabled in M, i.e. $Enbl(M) = \{t \in \mathcal{T} \mid Pre(\cdot, t) \leq M\}$.

A PN $\langle \mathcal{N}, M_0 \rangle$ is *1-safe* if for every place $p \in \mathcal{P}$ we have that $M(p) \leq 1$ for any marking M that is reachable from M_0.

Denote by $\mathcal{T}^*$ the Kleene closure of the set $\mathcal{T}$ and by ϵ the empty string. Then let $\sigma \in \mathcal{L}_{\mathcal{N}}(M_0) \subseteq \mathcal{T}^*$ and $\mathcal{T}' \subset \mathcal{T}$. The projection $\Pi_{\mathcal{T}'} : \mathcal{L}_{\mathcal{N}}(M_0) \to \mathcal{T}'^*$ is defined as: $i)$ $\Pi_{\mathcal{T}'}(\epsilon) = \epsilon$; $ii)$ $\Pi_{\mathcal{T}'}(t) = t$ if $t \in \mathcal{T}'$; $iii)$ $\Pi_{\mathcal{T}'}(t) = \epsilon$ if $t \in \mathcal{T} \setminus \mathcal{T}'$; and $iv)$ $\Pi_{\mathcal{T}'}(\sigma t) = \Pi_{\mathcal{T}'}(\sigma)\Pi_{\mathcal{T}'}(t)$ for $\sigma \in \mathcal{L}_{\mathcal{N}}(M_0)$ and $t \in \mathcal{T}$.

2.2. *Occurrence nets*

Definition 1. *Given a PN $\mathcal{N} = (\mathcal{P}, \mathcal{T}, F)$ the immediate dependence relation $\preceq_1 \subset (\mathcal{P} \times \mathcal{T}) \cup (\mathcal{T} \times \mathcal{P})$ is defined as:*

$$\forall (a,b) \in (\mathcal{P} \times \mathcal{T}) \cup (\mathcal{T} \times \mathcal{P}) : a \preceq_1 b \textit{ if } F(a,b) \neq 0$$

Define $\preceq$ as the transitive closure of $\preceq_1$ ($\preceq = \preceq_1^*$).

Definition 2. *Given a PN $\mathcal{N} = (\mathcal{P}, \mathcal{T}, F)$ the immediate conflict relation $\sharp_1 \subset \mathcal{T} \times \mathcal{T}$ is defined as:*

$$\forall (t_1, t_2) \in \mathcal{T} \times \mathcal{T} : t_1 \sharp_1 t_2 \textit{ if } {}^\bullet t_1 \cap {}^\bullet t_2 \neq \emptyset$$

Define $\sharp \subset (\mathcal{P} \cup \mathcal{T}) \times (\mathcal{P} \cup \mathcal{T})$ as $\forall (a,b) \in (\mathcal{P} \cup \mathcal{T}) \times (\mathcal{P} \cup \mathcal{T})$:

$$a \sharp b \textit{ if } \exists t_1, t_2 \textit{ s.t. } t_1 \sharp_1 t_2 \textit{ and } t_1 \preceq a \textit{ and } t_2 \preceq b$$

The independence relation $\| \subset (\mathcal{P} \cup \mathcal{T}) \times (\mathcal{P} \cup \mathcal{T})$ is defined as $\forall (a,b) \in (\mathcal{P} \cup \mathcal{T}) \times (\mathcal{P} \cup \mathcal{T})$:

$$a \| b \Rightarrow \neg(a \sharp b) \wedge (a \npreceq b) \wedge (b \npreceq a)$$

Definition 3. *Given two PNs $\mathcal{N} = (\mathcal{P}, \mathcal{T}, F)$ and $\mathcal{N}' = (\mathcal{P}', \mathcal{T}', F')$, ϕ is a homomorphism from $\mathcal{N}$ to $\mathcal{N}'$, denoted $\phi : \mathcal{N} \to \mathcal{N}'$ where:*

1) $\phi(\mathcal{P}) \subseteq \mathcal{P}'$ and $\phi(\mathcal{T}) \subseteq \mathcal{T}'$

2) $\forall t \in \mathcal{T}$, the restriction of ϕ to ${}^\bullet t$ is a bijection between ${}^\bullet t$ and ${}^\bullet \phi(t)$

3) $\forall t \in \mathcal{T}$, the restriction of ϕ to $t^\bullet$ is a bijection between $t^\bullet$ and $\phi(t)^\bullet$

Definition 4. *An occurrence net is a net $O = (B, E, \preceq_1)$ such that:*

i) $\forall a \in B \cup E : \neg(a \preceq a)$ (acyclic)

ii) $\forall a \in B \cup E : \mid \{b : a \preceq b\} \mid < \infty$ (well-formed)

iii) $\forall b \in B : \mid {}^\bullet b \mid \leq 1$ (no backward conflict)

In the following B is referred as the set of conditions while E is the set of events.

Definition 5. *A configuration $C = (B_C, E_C, \preceq)$ in the occurrence net O is defined as follows:*

i) C is a proper sub-net of O ($C \subseteq O$)

ii) C is conflict free, i.e. $\forall(a,b) \in (B_C \cup E_C) \times (B_C \cup E_C) \Rightarrow \neg(a\sharp b)$

iii) C is causally upward-closed, i.e. $\forall b \in B_C \cup E_C : a \in B \cup E$ and $a \preceq_1 b \Rightarrow a \in B_C \cup E_C$

iv) $\mathtt{min}_{\preceq}(C) = \mathtt{min}_{\preceq}(O)$

Definition 6. *Consider a PN $\langle \mathcal{N}, M_0 \rangle$ s.t. $\forall p \in \mathcal{P} : M_0(p) \in \{0,1\}$. A branching process $\mathcal{B}$ of a PN $\langle \mathcal{N}, M_0 \rangle$ is a pair $\mathcal{B} = (O, \phi)$ where O is an occurrence net and ϕ is a homomorphism $\phi : O \to \mathcal{N}$ s.t.:*

1) the restriction of ϕ to $\mathtt{min}_{\preceq}(O)$ is a bijection between $\mathtt{min}_{\preceq}(O)$ and M_0 (the set of initially marked places)

2) $\phi(B) \subseteq \mathcal{P}$ and $\phi(E) \subseteq \mathcal{T}$

3) $\forall a, b \in E$: (${}^\bullet a = {}^\bullet b$) $\wedge$ ($\phi(a) = \phi(b)$) $\Rightarrow a = b$

For a configuration C in O denote by $CUT(C)$ the maximal (w.r.t. set inclusion) set of conditions in C that have no successors in C:

$$CUT(C) = [(\bigcup_{e \in E_C} e^\bullet) \cup (\mathtt{min}_{\preceq}(O)] \setminus (\bigcup_{e \in E_C} {}^\bullet e)$$

Definition 7. *Given a PN $\langle \mathcal{N}, M_0 \rangle$ and two branching processes $\mathcal{B}, \mathcal{B}'$ of the PN $\langle \mathcal{N}, M_0 \rangle$ then $\mathcal{B}' \subseteq \mathcal{B}$ if there exists an injective homomorphism $\varphi : \mathcal{B}' \to \mathcal{B}$ s.t. $\varphi(\mathtt{min}(\mathcal{B}')) = \mathtt{min}(\mathcal{B})$ and $\phi \circ \varphi = \phi'$.*

There exists (up to an isomorphism) a unique maximum branching process (w.r.t. $\subseteq$) that is the unfolding of $\langle \mathcal{N}, M_0 \rangle$ and is denoted $\mathcal{U}_{\mathcal{N}}(M_0)$ [MCM 92].

Denote by $\mathcal{C}$ the set of all the configurations C of the unfolding $\mathcal{U}_{\mathcal{N}}(M_0)$. For a configuration $C \in \mathcal{C}$ denote by $\langle E_C \rangle_{\preceq}$ the set of strings that are linearizations of $(E_C, \preceq)$ where a string $\sigma = e_1 e_2 \dots e_\upsilon$ is a linearization of $(E_C, \preceq)$ if $\upsilon =\mid E_C \mid$ and $\forall e_\iota, e_\lambda \in E_C$ we have that: $i)$ $e_\iota = e_\lambda \Rightarrow \iota = \lambda$ and $ii)$ for $\iota \neq \lambda$, if $e_\iota \preceq e_\lambda$ then $\iota < \lambda$.

2.3. *Time Petri nets*

A Time Petri Net (TPN) $\mathcal{N}^\theta = (\mathcal{P}, \mathcal{T}, F, I^s)$, consists of an (untimed) Petri Net $\mathcal{N} = (\mathcal{P}, \mathcal{T}, F)$ (called the untimed support of $\mathcal{N}^\theta$) and the static time interval function $I^s : \mathcal{T} \to \mathcal{I}(\mathbb{Q}^+)$, $I^s(t) = [L^s_t, U^s_t]$, $L^s_t, U^s_t \in \mathbb{Q}^+$, representing the set of all possible time delays associated to transition $t \in \mathcal{T}$.

In a TPN $\langle \mathcal{N}^\theta, M^\theta_0 \rangle$ we say that a transition t becomes enabled at the time θ^{en}_t then the clock attached to t is started and the transition t can and must fire at some time $\theta_t \in [\theta^{en}_t + L^s_t, \theta^{en}_t + U^s_t]$, provided t did not become disabled because of the firing of another transition. Notice that t is forced to fire if it is still enabled at the time $\theta^{en}_t + U^s_t$.

Definition 8. *A state at the time θ (according to a global clock) of a TPN $\langle \mathcal{N}^\theta, M_0^\theta \rangle$ is a pair $S_\theta = (M, FI)$ where M is a marking and FI is a firing interval function associated to each enabled transition in M ($FI : \mathcal{T} \to \mathcal{I}(\mathbb{Q}^+)$).*

If t is executed at the time $\theta_t \in \mathbb{Q}^+$ we write $(M, FI) \xrightarrow{\langle t, \theta_t \rangle} (M', FI')$ or simply $S \xrightarrow{\langle t, \theta_t \rangle} S'$:

1) $(M \geq Pre(\cdot, t) \wedge \theta_t \geq \theta_t^{en} + L_t^s) \wedge (\forall t' \in \mathcal{T}$ s.t. $M \geq Pre(\cdot, t') \Rightarrow \theta_t \leq \theta_{t'}^{en} + U_{t'}^s)$

2) $M' = M - Pre(\cdot, t) + Post(t, \cdot)$

3) $\forall t'' \in \mathcal{T}$ s.t. $M' \geq Pre(\cdot, t'')$ we have:

- if $t'' \neq t \wedge M \geq Pre(\cdot, t'')$ then $FI(t'') = [\max(\theta_{t''}^{en} + L_{t''}^s, \theta_t), \theta_{t''}^{en} + U_{t''}^s]$

- else $\theta_{t''}^{en} = \theta_t$ and $FI(t'') = [\theta_{t''}^{en} + L_{t''}^s, \theta_{t''}^{en} + U_{t''}^s]$

A legal time trace τ^θ in a TPN $\mathcal{N}^\theta$ satisfies: $\tau^\theta = S_0 \xrightarrow{\langle t_1, \theta_{t_1} \rangle} S_1 \xrightarrow{\langle t_2, \theta_{t_2} \rangle} S_{v-1} \ldots \xrightarrow{\langle t_v, \theta_{t_v} \rangle} S_v$.

Definition 9. *Denote $\xrightarrow{*}$ the reflexive and transitive closure of $\to$. The state graph of a TPN $\mathcal{N}^\theta$ is $SG = (\mathcal{S}, \xrightarrow{*}, S_0)$ where $\mathcal{S} = \left\{ S \mid S_0 \xrightarrow{*} S \right\}$ is the set of reachable states from the initial state $S_0 = (M_0, FI_0)$ with $FI_0(t) = I^s(t)$ for all $t \in \mathcal{T}$ s.t. $M_0 \geq Pre(\cdot, t)$ otherwise $FI_0(t)$ is not defined.*

In the following for a time trace τ^θ we use the notation τ to denote its untimed support. For the initial state S_0 we use also the notation M_0^θ. Denote $\mathcal{L}_{\mathcal{N}^\theta}^\theta(M_0^\theta)$ the set of all legal time traces that can be executed in $\langle \mathcal{N}^\theta, M_0^\theta \rangle$. We call $\mathcal{L}_{\mathcal{N}^\theta}^\theta(M_0^\theta)$ the time language of the TPN $\langle \mathcal{N}^\theta, M_0^\theta \rangle$.

$\mathcal{L}_{\mathcal{N}^\theta}(M_0^\theta)$ is the untimed support language of the time language $\mathcal{L}_{\mathcal{N}^\theta}^\theta(M_0^\theta)$, i.e.:

$$\mathcal{L}_{\mathcal{N}^\theta}(M_0^\theta) = \left\{ \tau \mid \exists \tau^\theta \in \mathcal{L}_{\mathcal{N}^\theta}^\theta(M_0^\theta) \right\}$$

3. Diagnosis of TPNs

We consider the following plant description:

1) the TPN model $\langle \mathcal{N}^\theta, M_0^\theta \rangle$ is untimed 1-safe

2) $\mathcal{T} = \mathcal{T}_o \cup \mathcal{T}_{uo}$ where $\mathcal{T}_o$ is the set of observable events and $\mathcal{T}_{uo}$ is the set of unobservable (silent) events

3) l_o is the observation labeling function $l_o : \mathcal{T} \to \Omega_o \cup \{\epsilon\}$ where Ω_o is a set of labels and ϵ is the empty label. $l_o(t) = \epsilon$ if $t \in \mathcal{T}_{uo}$ and $l_o(t) \in \Omega_o$ if $t \in \mathcal{T}_o$

4) when an observable transition $t^o \in \mathcal{T}_o$ is executed in the plant the label $l_o(t^o)$ is emitted together with the global time $\theta_{l_o(t^o)}$ when this execution of t^o took place

5) the observation is always correct and the execution time of an observed event is measured with perfect accuracy according to a global clock, and received without delay

6) the execution of an unobservable event does not emit anything (is silent)

7) the faults are modeled by a subset of unobservable events, $\mathcal{T}_f \subseteq \mathcal{T}_{uo}$. $\mathcal{T}_f$ can be partitioned regarding the kinds of faults that may happen in the process as $\mathcal{T}_f = \mathcal{T}_{\mathtt{F}_1} \cup \mathcal{T}_{\mathtt{F}_2} \ldots \cup \mathcal{T}_{\mathtt{F}_m}$.

Formal description of the problem: *Given the plant model $\langle \mathcal{N}^\theta, M_0^\theta \rangle$ as described above, design an on-line algorithm that derives the fault diagnosis of the plant based on the model and the received observation. The exact meaning of diagnosis is defined below.*

We make the assumption that a cycle that contains only unobservable transitions that can be executed infinitely often contains at least one transition that has a non-zero lower bound of its static interval. This avoids the possibility of infinitely many events occurring at the same point in time.

The observation available to the diagnoser at the time the n^{th} observable event is executed in the plant is denoted as:

$$\mathcal{O}_n^\theta = \langle obs_1, \theta_{obs_1} \rangle, \ldots, \langle obs_n, \theta_{obs_n} \rangle$$

where $obs_1, \ldots, obs_n \in \Omega_o$ are the labels that are received and $\theta_{obs_1} \leq \theta_{obs_2} \ldots \leq \theta_{obs_n}$ are the times at which the corresponding events occur.

Denote by $\mathcal{O}_{n,\xi}^\theta$ the plant observation at the time $\xi > \theta_{obs_n}$, i.e. $\mathcal{O}_{n,\xi}^\theta$ includes $\mathcal{O}_n^\theta$ together with the information that no observation is received in the interval $[\theta_{obs_n}, \xi]$.

$\mathcal{L}_{\mathcal{N}^\theta}^\theta(M_0^\theta, \mathcal{O}_n^\theta)$ is the set of all time traces that are feasible in $\langle \mathcal{N}^\theta, M_0^\theta \rangle$ up to the time of the last observation θ_{obs_n} and that obey the received observation $\mathcal{O}_n^\theta$ where $\tau^\theta \in \mathcal{L}_{\mathcal{N}^\theta}^\theta(M_0^\theta, \mathcal{O}_n^\theta)$ if:

1) $\tau^\theta \in \mathcal{L}_{\mathcal{N}^\theta}^\theta(M_0^\theta, \theta_{obs_n})$ (τ^θ is legal)

2) $l_o(\tau) = obs_1, \ldots, obs_n$ (τ^θ obeys the *"untimed"* observation)

3) for each observable transition $t_k^o \in \mathcal{T}_o$, $k = 1, \ldots, n$ we have that $l_o(t_k^o) = obs_k \Rightarrow \theta_{t_k} = obs_k$ (τ^θ obeys the execution times of the observed transitions)

where $\tau^\theta \in \mathcal{L}_{\mathcal{N}^\theta}^\theta(M_0^\theta, \theta_{obs_n})$ if the execution time of the last event in τ^θ is smaller than or equal θ_{obs_n}.

Similarly $\mathcal{L}_{\mathcal{N}^\theta}^\theta(M_0^\theta, \mathcal{O}_{n,\xi}^\theta)$ is the set of all time traces that are feasible in $\langle \mathcal{N}^\theta, M_0^\theta \rangle$ up to the time ξ and that obey the received observation $\mathcal{O}_{n,\xi}^\theta$.

The plant diagnosis $\mathcal{D}_{\mathcal{N}^\theta}(\mathcal{O}_{n,\xi}^\theta)$ based on the received observation $\mathcal{O}_{n,\xi}^\theta$ comprises the untimed strings obtained by projecting the untimed support traces contained in $\mathcal{L}_{\mathcal{N}^\theta}(M_0^\theta, \mathcal{O}_{n,\xi}^\theta)$ onto the set of fault transitions $\mathcal{T}_f$:

$$\mathcal{D}_{\mathcal{N}^\theta}(\mathcal{O}_{n,\xi}^\theta) = \left\{ \tau_f \mid \tau^\theta \in \mathcal{L}_{\mathcal{N}^\theta}^\theta(M_0^\theta, \mathcal{O}_{n,\xi}^\theta) \text{ and } \tau_f = \Pi_{\mathcal{T}_f}(\tau) \right\} \quad (1)$$

The diagnosis result of the plant at time ξ with the received observation $\mathcal{O}^{\theta}_{n,\xi}$ is:

$$\mathcal{DR}_{\mathcal{N}^{\theta}}(\mathcal{O}^{\theta}_{n,\xi}) = \begin{cases} \mathtt{F} \text{ iff } \epsilon \notin \mathcal{D}_{\mathcal{N}^{\theta}}(\mathcal{O}^{\theta}_{n,\xi}) \\ \mathtt{N} \text{ iff } \{\epsilon\} = \mathcal{D}_{\mathcal{N}^{\theta}}(\mathcal{O}^{\theta}_{n,\xi}) \\ \mathtt{UF} \text{ otherwise} \end{cases} \tag{2}$$

where similarly as defined for the untimed case in [SAM 95] we have that:

1) $\mathtt{F}$ means that a fault did necessarily happen in the plant:

$$\forall \tau^{\theta}_{f} \in \mathcal{D}_{\mathcal{N}^{\theta}}(\mathcal{O}^{\theta}_{n,\xi}) : \Pi_{\mathcal{T}_f}(\tau_f) \neq \epsilon$$

2) $\mathtt{N}$ means that a fault did not happen in the plant:

$$\forall \tau^{\theta}_{f} \in \mathcal{D}_{\mathcal{N}^{\theta}}(\mathcal{O}^{\theta}_{n,\xi}) : \Pi_{\mathcal{T}_f}(\tau_f) = \epsilon$$

3) $\mathtt{UF}$ means that it is uncertain whether a fault happened or not in the plant that is, there exist two legal time-strings $\tau^{\theta}_{f}, \tau'^{\theta}_{f} \in \mathcal{D}_{\mathcal{N}^{\theta}}(\mathcal{O}^{\theta}_{n,\xi})$ s.t. $\Pi_{\mathcal{T}_f}(\tau_f) \neq \epsilon$ and $\Pi_{\mathcal{T}_f}(\tau'_f) = \epsilon$.

In order to address properly the on-line fault diagnosis problem one must assume that the faults are unpredictable, i.e. faults cannot be detected that will happen for sure in the future. Otherwise one should be required to make calculations in advance so as to detect the imminent occurrence of a fault at the earliest time possible.

We illustrate this via the following example.

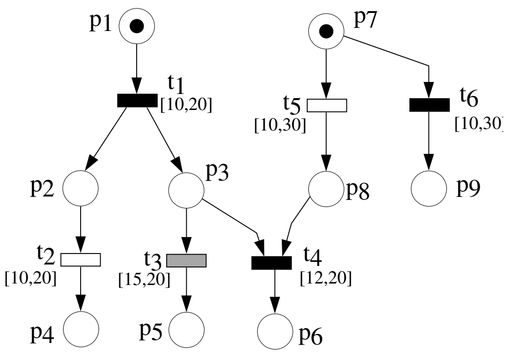

Figure 1. *The TPN of Example 1.*

Example 1. *Consider the TPN displayed in Fig.1. Static intervals are attached to each transition. t_2 and t_5 are observable transitions and t_3 is a fault transition.*

Consider that the occurrence of t_2 is observed at the (global) time 20. We claim that having this observation we know that the fault event t_3 happens for sure in the future. This is because t_1 happened at the time $\theta_{t_1} = 10$ that means that the fault *transition t_3 is forced to fire at the time 30. Since the observable transition t_2 was not executed yet, it means that t_4 can become enabled only after the time 20 and can fire at the earliest time 32 that is after the time t_3 is forced to fire.*

Thus after the first observation and regardless of the time of the next observation, it is certain that a fault will happen. This means that the plant analysis should be developed up to an arbitrary large time in the future, and to check each time if a fault becomes imminent in the future given the plant observation up to the current time of the process. But this is very inefficient since it is known that the state space of TPNs of reasonable size can be very large.

In order to address properly the problem of on-line fault diagnosis we assume that the faults are unpredictable, i.e. given any observation generated by the plant one cannot predict that a fault will happen for sure in the future. Unfortunately, this condition cannot be checked for general TPNs unless expensive calculations on the complete state space of the model are made.

However we impose a structural assumption on the TPN model that is a sufficient condition for the faults to be unpredictable. It simply says that for any fault transition $t_f \in \mathcal{T}_f$, there is a non-fault transition t that has its pre-set ${}^\bullet t$ included in or equal to the pre-set of t_f (${}^\bullet t \subseteq {}^\bullet t_f$) and moreover the lower bound of the static interval of the normal transition L_t^s is not larger than the upper bound of the static interval of the fault transition t_f ($L_t^s \leq U_t^s$).

Assumption 1. *The TPN model $\mathcal{N}^\theta = (\mathcal{P}, \mathcal{T}, F, I^s)$ is such that $\forall t \in \mathcal{T}_f$, $\exists t' \in \mathcal{T} \backslash \mathcal{T}_f$ s.t. $i)$ ${}^\bullet t' \subseteq^\bullet t$ and $ii)$ $L_{t'}^s \leq U_t^s$.*

It is easy to see now that if Assumption 1 is satisfied then the faults are unpredictable. This is because the time a fault transition t_f becomes enabled is not smaller than the time the normal transition t becomes enabled (condition $i)$) and t_f is not forced to fire at a time before t can fire (condition $ii)$).

4. The analysis of TPNs based on partial orders

In this section we present the analysis of TPNs based on partial orders. The reason is that the state class graph methods [BER 83],[YON 98] have the drawback, when applied for partial observable TPNs as proposed in [GHA 05], that they consider all the interleavings of the unobservable transitions. Even though not all the interleavings of the untimed concurrent transitions can be possible in a TPN, their consideration makes the analysis of TPNs of reasonable size sometimes impossible.

Example 2. *Consider the TPN displayed in Fig. 2. Static intervals are attached to each transition. The observable transitions are t_4, t_7 and t_{10} and they emit the same label. t_3 and t_9 are faulty transitions.*

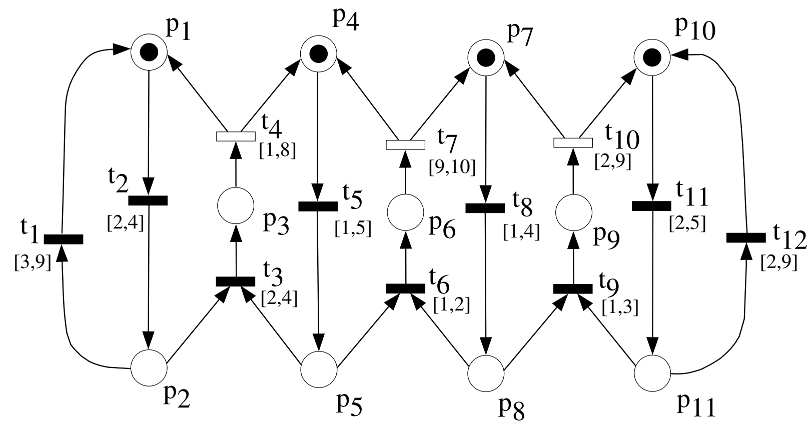

Figure 2. *The TPN of the Example 2.*

For instance if the plant analysis is based on the state class graph construction [BER 83] one should consider when the process starts all the possible interleavings of the unobservable concurrent transitions t_2, t_5, t_8, and t_{11}.

Hence in this example the timing information does not reduce the number of the interleavings of the concurrent (unobservable) events that are considered. The partial order reduction techniques developed for untimed PN [MCM 92],[ESP 94], [BEN 03] are shown in [HUL 95], [SEM 96], [AUR 97], [CHA 05] to be applicable for TPN. Consider a configuration C in the unfolding $\mathcal{U}_{\mathcal{N}}(M_0)$ of the untimed PN support of a TPN. Then consider a valuation Θ of the execution times at which the events $e \in E_C$ in the configuration C are executed. I.e. for each $e \in E_C$ consider a time value $\theta_e \in \mathbb{T}$ ($\mathbb{T}$ the time axis) at which e occurs and Θ is an $\mid E_C \mid$-tuple comprising all the values at which all the events $e \in E_C$ are executed.

An untimed configuration C together with a valuation $\Theta \in \mathbb{T}^{|E_C|}$ of the execution time for its events is called a time configuration (time process in [AUR 97]) of the TPN model.

A time configuration is legal if there is a legal trace $\tau^\theta \in \mathcal{L}^\theta_{\mathcal{N}^\theta}(M^\theta_0)$ in the TPN $\langle \mathcal{N}^\theta, M^\theta_0 \rangle$ whose untimed support τ is a linearization of the partial order relation of the events in the configuration (i.e. $\tau = \phi(\sigma)$ and $\sigma \in \langle E_C \rangle_{\preceq}$) while the execution time θ_t of every transition t considered in the trace τ^θ is identical with the valuation θ_e of the event e for which t is its image via ϕ.

Consider an untimed configuration $C \in \mathcal{C}$. The TPN C^θ is obtained from the untimed configuration C attaching to each event the static interval I^s_t that corresponds in the original TPN to transition t s.t. $\phi(e) = t$.

$$C^\theta = (B_C, E_C, \preceq, \mathtt{min}_{\preceq}(\mathcal{U}_{\mathcal{N}}), I^s)$$

where:

1) B_C is the set of places

2) E_C is the set of events (transitions)

3) $\preceq$ is the incidence function

4) $\mathtt{min}_{\preceq}(\mathcal{U}_{\mathcal{N}})$ is the initial marking (the tokens "arrive" in these places at the time when the process starts)

5) $I^s : E_C \to \mathcal{I}(\mathbb{T}_+)$, $I^s(e) = I^s(t)$ with $t = \phi(e)$

Denote by $\widetilde{K}_{C^\theta}$ the following system of inequalities:

$$\widetilde{K}_{C^\theta} = \Big\{ \max_{e' \in {}^{\bullet\bullet}e} (\theta_{e'}) + L^s_e \leq \theta_e \leq \max_{e' \in {}^{\bullet\bullet}e} (\theta_{e'}) + U^s_e \quad \text{for all} \;\; e \in E_C \tag{3}$$

where in (3) ${}^{\bullet\bullet}e = \emptyset$ implies $\max_{e' \in {}^{\bullet\bullet}e}(\theta_{e'}) = 0$.

Proposition 1. $\forall \tau^\theta \in \mathcal{L}^\theta_{\mathcal{N}^\theta}(M^\theta_0)$ *we have that if* $\tau = \phi(\sigma)$ *and* $\sigma \in \langle E_C \rangle_{\preceq}$*, then* Θ *is a solution of* $\widetilde{K}_{C^\theta}$*, where* $\Theta = (\theta_{t_1}, \ldots, \theta_{t_{|E_C|}}) = (\theta_{e_1}, \ldots, \theta_{e_{|E_C|}})$ *with* $\phi(e_i) = t_i$*,* $i = 1, \ldots, \mid E_C \mid$.

Proof. The proof is straightforward since for 1-safe PN there exists an unique configuration C in the net unfolding $\mathcal{U}_{\mathcal{N}}$ s.t. $\tau = \phi(\sigma)$ and $\sigma \in \langle E_C \rangle_{\preceq}$. Obviously the conditions required for Θ to be a solution of $\widetilde{K}_{C^\theta}$ are satisfied by any legal time trace. □

Denote by $Sol(\widetilde{K}_{C^\theta})$ the set of all solutions of $\widetilde{K}_{C^\theta}$. The $\mid E_C \mid$-hyperbox $\widetilde{\mathbf{I}}$ that circumscribes $Sol(\widetilde{K}_{C^\theta})$ is easily obtained in the following way:

1) $\forall e \in E_C$ s.t. ${}^{\bullet\bullet}e = \emptyset$, $\widetilde{I}(e) = [\widetilde{L}(e), \widetilde{U}(e)]$ with $\widetilde{L}(e) = L^s_e$ and $U(e) = U^s_e$

2) $e \in E_C$ s.t. ${}^{\bullet\bullet}e \neq \emptyset$, $\widetilde{I}(e) = [\widetilde{L}(e), \widetilde{U}(e)]$ with $\widetilde{L}(e) = \max_{e' \in {}^{\bullet\bullet}e}(\widetilde{L}(e')) + L^s_e$ and $\widetilde{U}(e) = \max_{e' \in {}^{\bullet\bullet}e}(\widetilde{U}(e')) + U^s_e$

We cannot claim yet that for $\forall C \in \mathcal{C}$ there exists at least a legal time configuration that corresponds with C because for a general TPN the enabling of a transition does not guarantee that it eventually fires because some conflicting transition may be forced to fire before.

Denote by $\breve{E}_C$ the set of conflicting events of a configuration $C \in \mathcal{C}$ where E_C comprises the events that could have been executed but are not included in E_C:

$$\breve{E}_C = \{\breve{e} \in E \setminus E_C \mid {}^{\bullet}\breve{e} \subseteq B_C\}$$

The *characteristic system* K_{C^θ} of configuration $C^\theta \in \mathcal{C}$ is obtained adding to $\widetilde{K}_{C^\theta}$ all the inequalities regarding all the conflicting events :

$$K_{C^\theta} = \begin{cases} \max\limits_{e' \in {}^{\bullet\bullet}e} (\theta_{e'}) + L^s_e \leq \; \theta_e \leq \max\limits_{e' \in {}^{\bullet\bullet}e} (\theta_{e'}) + U^s_e \;\; \text{for all} \;\; e \in E_C \\[2ex] \min\limits_{e' \sharp_1 \breve{e}} (\theta_{e'}) \leq \max\limits_{e'' \in \, {}^{\bullet\bullet}\breve{e}} (\theta_{e''}) + U^s_{\breve{e}} \;\; \text{for all} \;\; \breve{e} \in \breve{E}_C \end{cases} \tag{4}$$

Proposition 2. *Given an arbitrary time ξ we have that $\tau^\theta \in \mathcal{L}^\theta_{\mathcal{N}^\theta}(M^\theta_0, \xi)$ if:*

1) $\tau = \phi(\sigma)$, $\sigma \in \langle E_C \rangle_\preceq$ and $C \in \mathcal{C}$

2) Θ is a solution of K_{C^θ}

3) $\forall e \in E_C \Rightarrow \theta_e \leq \xi$,

4) $\forall e \in Enbl(C)$, $\max_{e' \in {}^{\bullet\bullet}e}(\theta_{e'}) + U^s_e \geq \xi$.

Proof. $\Rightarrow$ Condition 1 and 3 are trivial and the proof that $\Theta = (t_1, \ldots, t_n)$ is a solution of K_{C^θ} is by induction.
$\Leftarrow$ The proof is trivial. □

The problem the we should answer next is:

Up to what time ξ to make the calculations for the on-line monitoring ?

There are different solutions to answer this question, depending on the computational capability, the plant behavior, and the requirements for the diagnosis result.

Solution 1: Calculations in advance

This solution is appropriate for a plant known to have a cyclic behavior, such that periodically the plant halts in a *"quiescent state"* [BAR 99] (a state s.t. no transition is executed until a next trigger-event is executed). E.g. the protection system in an electrical network is triggered by the occurrence of a short-circuit and the plant returns to a quiescent state after the fault is cleared. Another example is the cyclic operation of a plant, where each operation cycle is initiated by the plant operator.

Having derived the plant behavior up to some time $\widehat{\xi}$, the plant is monitored on-line in the following way:

1) the received observation is taken into account by adding (in)equality constraints to the characteristic system of a configuration

2) or discarding configurations when the current time exceeds the latest execution time of an observable event in a configuration.

Remark 1. *The faults are considered under Assumption 1. Thus only fault transitions that are executed before the current time of the plant can be detected to have happened for sure. In other words, computing the future plant behavior does not change the detection of the faults w.r.t. the state* `F` *of the diagnoser.*

The main drawback of this method is that a large amount of calculations is required to be performed in advance and then discarded because of the received observation.

Solution 2: Calculations after each observation

The second solution is to perform calculations each time an event is observed in the plant. E.g. when the first observable event is executed in the plant we derive the plant behavior up to the time θ_{obs_1} in the following way.

Let the first observation be $\mathcal{O}_1^\theta = \langle obs_1, \theta_{obs_1} \rangle$. Consider the set of configurations $\mathcal{C}(\mathcal{O}_1^\theta)$ s.t. $C \in \mathcal{C}(\mathcal{O}_1^\theta)$ if:

1) E_C contains only one event e^o that corresponds with an observable event

2) $\phi(e^o) = t^o$ and $\ell(t^o) = obs_1$ and $\theta_{obs_1} \in \breve{I}(e^o)$

3) $\forall e \in {}^\bullet CUT(C) \Rightarrow \widetilde{L}(e) < \theta_{obs_1}$

4) $\forall e \in Enbl(C) \Rightarrow \widetilde{U}(e) > \theta_{obs_1}$

where $Enbl(C)$ denotes the set of events that correspond via ϕ to transitions that are enabled from the marking $\phi(CUT(C))$.

The characteristic system $K_{C^\theta}(\mathcal{O}_1^\theta)$ of configuration $C^\theta \in \mathcal{C}(\mathcal{O}_1^\theta)$ is obtained adding to $\widetilde{K}_{C^\theta}$ inequalities regarding the conflicting events and the received observation:

$$K_{C^\theta}(\mathcal{O}_1^\theta) = \begin{cases} \max\limits_{e' \in {}^{\bullet\bullet}e}(\theta_{e'}) + L_e^s \leq \theta_e \leq \max\limits_{e' \in {}^{\bullet\bullet}e}(\theta_{e'}) + U_e^s \text{ for all } e \in E_C \\ \min\limits_{e' \sharp_1 \breve{e}}(\theta_{e'}) \leq \max\limits_{e'' \in {}^{\bullet\bullet}\breve{e}}(\theta_{e''}) + U_{\breve{e}}^s \text{ for all } \breve{e} \in \breve{E}_C \\ \theta_{e^o} = \theta_{obs_1} \text{ for } \phi(e^o) = t^o \wedge \ell(t^o) = obs_1 \\ \theta_{e'^o} \geq \theta_{obs_1} \text{ for all } e'^o \in Enbl(C) \end{cases} \tag{5}$$

This method requires less computations but the price to be paid is that a fault may be detected with a delay. This is because no calculations are performed until a new observation is received, thus the fact that the current time of the plant exceeds the latest execution time of an observable event is not taken into account.

However this method can be applied when the rate of receiving observations is high, i.e. the time interval in between two observations is short and control actions are taken with some latency.

Solution 3: Calculations up to a discarding time

A discarding time is the earliest time when in absence of any observation one can discard untimed support traces because it can be proved that they are not valid. E.g. the first discarding time is the smallest among the latest execution times of an observable transition in the plant.

Definition 10. *A configuration $C_v \in \mathcal{C}$ is generated up to the time ξ if:*

1) $\max_{e \in {}^{\bullet}CUT(C_\upsilon)}(\widetilde{L}_\upsilon(e)) \leq \xi$

2) $\min_{e \in Enbl(C_\upsilon)}(\widetilde{U}_\upsilon(e)) > \xi$

Given a configuration $C_\nu \in \mathcal{C}$ that is generated up to a time ξ', denote by $\mathcal{C}_\nu(\xi)$ the set of extensions of C_ν up to the time $\xi > \xi'$ where $C_{\ell_\nu} \in \mathcal{C}_\nu(\xi)$ if:

– $C_\nu \subseteq C_{\ell_\nu}$ (C_{ℓ_ν} is a continuation of C_ν)

– and C_{ℓ_ν} is generated up to the time ξ.

The first discarding time $\hat{\theta}$ is calculated iteratively as follows.

Starting from the initial configuration $C^\perp = (B^\perp, E^\perp, \preceq_1)$ we construct an initial part of the net unfolding by appending events as in the untimed case, the only difference being that among all the enabled events only the events with the smallest upper bound $\widetilde{U}(e)$ are appended, until the first observable event say e^o is encountered.

The discarding time is set equal with $\widetilde{U}(e^o)$ and then the configurations that contain e^o are extended up to the time $\widetilde{U}(e^o)$. Denote this set by $\mathcal{C}^{new}_{obs}$. Then for each configuration $C_\nu \in \mathcal{C}^{new}_{obs}$ we calculate $Sol(K_{C^\theta_\nu})$ and for those configurations that have a non-empty solution set we calculate $U_\nu(e'^o)$, i.e. the latest time when e^o can be executed. Obviously $U_\nu(e'^o) \leq \widetilde{U}_\nu(e^o)$.

The discarding time $\hat{\theta}$ is set as the smallest among the latest times when an observable event can be executed considering all $C_\nu \in \mathcal{C}^{new}_{obs}$. Notice that a configuration C_ν may contain some other observable events and after calculating $Sol(K_{C^\theta_\nu})$ some other observable event may have the minimal latest time for its execution.

Recursively all the configurations that contain only unobservable events are extended up to the new discarding time $\hat{\theta}$ by appending those events among all the enabled events with the smallest upper bound $\widetilde{U}(e)$ until either a new observable event is encountered or no more events can be appended.

Algorithm 1 provides the pseudo-code for the computation of the first discarding time. Notice that because $\hat{\theta}$ is calculated recursively some configurations (that contain at least one observable event) are generated up to times bigger than $\hat{\theta}$. However this does not affect the diagnosis result since the events that can be executed after the time $\hat{\theta}$ are seen as a prognosis.

Example 3. *Consider the TPN displayed in Fig. 2. The first discarding time is calculated as follows. (see Fig. 3 where a part of the unfolding $\mathcal{U}_\mathcal{N}(M_0)$ is drawn attaching to each event e $\in E$ the time interval $\widetilde{I}(e)$).*

1) first e_2 and e_8 are appended since they have the smallest upper bound among the enabled events from the initial marking ($\widetilde{U}(e_2) = \widetilde{U}(e_8) = 4$)

2) then e_5 and e_{11} are appended

3) unobservable transitions are appended until a first observable transition is appended; in this example this can be either e_4, e_7 or e_{10}

Algorithm 1 Discarding_time

```
Require: $C^{\perp} = (B^{\perp}, E^{\perp}, \preceq_1)$
Ensure: $\hat{\theta}$
$All\mathcal{C} = \left\{C^{\perp}\right\}, \hat{\theta} = +\infty$
repeat
    $Enbl(All\mathcal{C}) = \bigcup_{C \in All\mathcal{C}} Enbl(C)$
    $APP(All\mathcal{C}) = \left\{e \in Enbl(All\mathcal{C}) \mid \forall e' \in Enbl, \widetilde{U}(e) \leq \widetilde{U}(e') \wedge \widetilde{L}(e) \leq \hat{\theta}\right\}$
    $APP^{obs}(All\mathcal{C}) = \{e \in APP(All\mathcal{C}) \mid \phi(e) \in \mathcal{T}_o\}$
    if $APP^{obs}(All\mathcal{C}) \neq \emptyset$ then
        $\mathcal{C}^{new}_{obs} = \left\{C^{new}_{obs} \mid C \in All\mathcal{C}, e^o \in APP^{obs}(All\mathcal{C}), C^{new}_{obs} = C \odot e^o\right\}$
        for all $C_\nu \in \mathcal{C}^{new}_{obs}$ do
            calculate $\mathcal{C}_\nu(\hat{\theta}_\nu)$ as the set of extensions of $C_\nu$ up to the time $\hat{\theta}_\nu = \min(\widetilde{U}(e^o), \hat{\theta})$
            for all $C_{\ell_\nu} \in \mathcal{C}_\nu(\hat{\theta}_\nu)$ do
                calculate $Sol(K_{C_{\ell_\nu}})$
                if $Sol(K_{C_{\ell_\nu}}) \neq \emptyset$ then
                    calculate the smallest $U_{\ell_\nu}(e'^o)$ for $e'^o \in E_{C_{\ell_\nu}}, \phi(e'^o) \in \mathcal{T}_o$
                    $\mathcal{C}_{obs} = \mathcal{C}_{obs} \cup C_{\ell_\nu}$
                end if
            end for
            $\hat{\theta}_\nu = \min_{C_{\ell_\nu} \in \mathcal{C}_\nu(\hat{\theta}_\nu)}(U_{\ell_\nu}(e^o))$
        end for
        $\hat{\theta} = \min(\hat{\theta}, \min_{C_\nu \in \mathcal{C}^{new}_{obs}}(\hat{\theta}_\nu))$
    else
        $\mathcal{C}^{new}_{uno} = \{C^{new}_{uno} \mid C \in All\mathcal{C}, e \in APP(All\mathcal{C}), C^{new}_{uno} = C \odot e\}$
        $All\mathcal{C} = All\mathcal{C} \cup \mathcal{C}^{new}_{uno}$
    end if
until $APP(All\mathcal{C}) = \emptyset$
```

4) then transitions are appended until the enabled transitions have their lower bound bigger than $\hat{\theta}$

5) for each configuration that contains an event say e that corresponds with $\hat{\theta}$ we calculate the solution set

6) if calculating the solution set we obtain for an observable event e that $U(e) < \hat{\theta}$, then $U(e)$ becomes the new discarding time.

The on-line diagnosis algorithm works as follows. When the process starts we derive the set of configurations up to the first discarding time and then we have two cases:

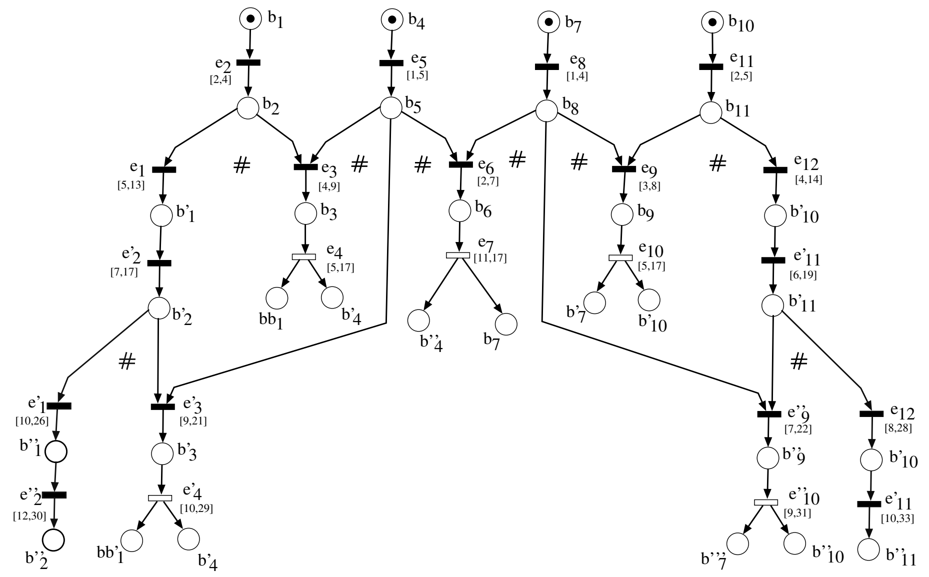

Figure 3. *A part of the unfolding of the TPN displayed in Fig. 2.*

Case 1 If no observation is received until the time of the process becomes equal with the discarding time $\hat{\theta}$ then:

1) the configurations that contain observable events with the upper bound corresponding with $\hat{\theta}$ are discarded

2) for all the other configurations that contain observable events inequalities of the form:

$$\mathcal{K}_{obs_1} = \left\{ \theta_{e^o} > \hat{\theta} \mid e^o \in E_C \text{ and } \phi(e^o) \in \mathcal{T}_o \right\}$$

are added to the characteristic systems K_{C^θ} and we derive the entire solution set

3) for all the configurations $C_v \in \mathcal{C}_{uno}$ that contain only unobservable events we check only if $Sol(K_{C_v^\theta})$ has a non-empty set of solutions

4) denote by $\mathcal{E}(\mathcal{O}_1^\theta)$ the set of traces that are obtained as linearizations of the set of events of the configurations that are not discarded

5) the diagnosis result $\mathcal{D}_{\mathcal{N}^\theta}^{po}(\mathcal{O}_{0,\hat{\theta}}^\theta)$ is obtained by projecting $\mathcal{E}(\mathcal{O}_1^\theta)$ onto $\mathcal{T}_f$.

Case 2 If the first observation $\langle obs_1, \theta_{obs_1} \rangle$ is received before the time of the process becomes equal with the discarding time $\hat{\theta}$ then:

1) the set of configurations $\mathcal{C}_{uno}$ that contain only unobservable events is discarded

2) for all the other configurations $C_\nu \in \mathcal{C}_{obs}$ that contain observable events an equality relation:

$$\mathcal{K}'_{obs_1} = \{\theta_{e^o} = \theta_{obs_1} \mid l_o(e^o) = obs_1 \wedge e^o \in C_\nu\}$$

and for observable events other than e^o inequalities of the form:

$$\mathcal{K}''_{obs_1} = \left\{\theta_{e'^o} > \hat{\theta} \mid e'^o \in E_C \text{ and } \phi(e'^o) \in \mathcal{T}_o\right\}$$

are added to the characteristic systems K_{C^θ} and then we derive the entire solution set

3) denote by $\mathcal{E}(\mathcal{O}_1^\theta)$ the set of traces that are obtained as linearizations of the set of events of the configurations that are not discarded

4) the diagnosis result $\mathcal{D}^{po}_{\mathcal{N}^\theta}(\mathcal{O}_1^\theta)$ is obtained by projecting $\mathcal{E}(\mathcal{O}_1^\theta)$ onto $\mathcal{T}_f$.

Algorithm 2 provides the pseudo-code for the plant diagnosis based on partial orders.

Algorithm 2 Diagnosis_1

```
Require: $\langle \mathcal{N}^\theta, M_0^\theta \rangle, \mathcal{T}_o, \mathcal{T}_{uo}$
Ensure: $\mathcal{DR}_{\mathcal{N}^\theta}(\mathcal{O}_\xi^\theta)$
Discarding_time($C^\perp$)
if $\theta_{obs_1} < \hat{\theta}$ then
    for all $C_\nu \in \mathcal{C}_{obs}$ s.t. $\exists e \in E_{C_\nu}, l_o(e) = obs_1$ do
        calculate $Sol(K_{C_\nu^\theta} \wedge \mathcal{K}'_{obs_1} \wedge \mathcal{K}''_{obs_1})$
        if $Sol(K_{C_\nu^\theta} \wedge \mathcal{K}_{obs_1}) \neq \emptyset$ then
            $\mathcal{E}(\mathcal{O}_1^\theta) = \mathcal{E}(\mathcal{O}_1^\theta) \cup \{\sigma \mid \sigma \in \langle E_{C_\nu} \rangle_\preceq\}$
        end if
    end for
    $\mathcal{D}_{\mathcal{N}^\theta}(\mathcal{O}_1^\theta) = \left\{\tau_f \mid \tau_f = \Pi_{\mathcal{T}_f}\tau \wedge \tau = \phi(\sigma) \wedge \sigma \in \mathcal{E}(\mathcal{O}_1^\theta)\right\}$
else
    for all $C_\upsilon \in \mathcal{C}_{obs} \cup \mathcal{C}_{uno}$ do
        check $Sol(K_{C_\upsilon^\theta} \wedge \mathcal{K}_{obs_1})$
        if $Sol(K_{C_\nu^\theta} \wedge \mathcal{K}_{obs_1}) \neq \emptyset$ then
            $\mathcal{E}(\mathcal{O}_{0,\hat{\theta}}^\theta) = \mathcal{E}(\mathcal{O}_1^\theta) \cup \{\sigma \mid \sigma \in \langle E_{C_\nu} \rangle_\preceq\}$
        end if
    end for
    $\mathcal{D}^{po}_{\mathcal{N}^\theta}(\mathcal{O}_{0,\hat{\theta}}^\theta) = \left\{\tau_f \mid \tau_f = \Pi_{\mathcal{T}_f}\tau \wedge \tau = \phi(\sigma) \wedge \sigma \in \mathcal{E}(\mathcal{O}_{0,\hat{\theta}}^\theta)\right\}$
end if
```

Notice that the plant diagnosis is derived either at the time θ_{obs_1} of the first observed event $\mathcal{D}^{po}_{\mathcal{N}^\theta}(\mathcal{O}_1^\theta)$ or in absence of any observation at the first discarding time $\hat{\theta}$ $\mathcal{D}^{po}_{\mathcal{N}^\theta}(\mathcal{O}_{0,\hat{\theta}}^\theta)$. We have that:

Theorem 1. *Given a TPN model $\langle \mathcal{N}^\theta, M_0^\theta \rangle$ we have that:*

1) when the first observable event is executed:

$$\mathcal{DR}_{\mathcal{N}^\theta}(\mathcal{O}_1^\theta) = \{\mathtt{F}\} \Leftrightarrow \mathcal{DR}^{po}_{\mathcal{N}^\theta}(\mathcal{O}_1^\theta) = \{\mathtt{F}\}$$

2) if no observation is received until the first discarding $\hat{\theta}$:

$$\mathcal{DR}_{\mathcal{N}^\theta}(\mathcal{O}_{0,\hat{\theta}}^\theta) = \{\mathtt{F}\} \Leftrightarrow \mathcal{DR}^{po}_{\mathcal{N}^\theta}(\mathcal{O}_{0,\hat{\theta}}^\theta) = \{\mathtt{F}\}$$

3) and for any time $\xi \leq \hat{\theta}$, in absence of any observation, the diagnosis result is different from F*:*

$$\mathcal{DR}_{\mathcal{N}^\theta}(\mathcal{O}_{0,\xi}^\theta) \neq \{\mathtt{F}\}$$

Proof. (1) and (2) have a similar proof. Based on Proposition 1 we calculate the set of legal traces up to a given time ξ. However some configurations include events that are executed after the time θ_{obs_1} or $\hat{\theta}$. Since the faults are unpredictable the consideration of some events that can be executed after the time θ_{obs_1} or $\hat{\theta}$ does not change the diagnosis result w.r.t. the detection of faults that for sure happened. (3) is proved straightforwardly by the assumption that the faults are unpredictable. □

Remark 2. *Obviously by imposing the inequalities that all the events in a configuration have execution times smaller than θ_{obs_1} or $\hat{\theta}$ one can derive exactly the diagnosis result at the time θ_{obs_1} respectively $\hat{\theta}$. However this is not efficient for practical calculations especially when the frequency of observations is high. Notice also that calculations in advance are not fully developed, thus it may be that an event that is considered executed after θ_{obs_1} might not be executed since an event that is successor of the observed event can pre-empt its execution.*

In what follows we present two methods to derive the solution set of the characteristic system of a configuration. The first method is based on the Extended Linear Complementarity Problem and derives the entire solution set as a union of faces of a polyhedron that satisfy the cross-complementarity condition [DES 95].

The second method is based on constraint propagation and derives for a configuration C a set of $\mid E_C \mid$-hyperboxes s.t. the union of the subsets of solutions that are circumscribed by the $\mid E_C \mid$-hyperboxes is a cover of the entire solution set.

5. The Extended Linear Complementarity Problem

The ELCP is defined as follows (see [DES 95]):

Given $A \in \mathbb{R}^{w\times z}$, $G \in \mathbb{R}^{q\times z}$, $c \in \mathbb{R}^{w}$, $d \in \mathbb{R}^{q}$, and m index sets $\psi_1, \ldots, \psi_m \subseteq \{1, \ldots, w\}$, find $x \in \mathbb{R}^{z}$ such that

$$Ax \geq c, \qquad Gx = d \tag{6}$$

$$\sum_{j=1}^{m} \prod_{i\in\psi_j} (Ax - c)_i \;=\; 0 \;\; . \tag{7}$$

Condition (7) can be interpreted as follows. Since $Ax \geq c$, all the terms in (7) are nonnegative. Hence, (7) is equivalent to $\prod_{i\in\psi_j}(Ax - c)_i = 0$ for $j = 1, \ldots, m$. So we could say that each set ψ_j corresponds to a group of inequalities in $Ax \geq c$, and that in each group at least one inequality should hold with equality. In [DES 95] an algorithm to find *all* solutions of an ELCP was developed. This algorithm yields a description of the complete solution set of an ELCP by finite points, generators for extreme rays, and a basis for the linear subspace associated with the maximal affine subspace of the solution set of the ELCP.

Let us now explain how $(max, +)$ equations of the form

$$\max_{i\in\mathcal{J}}(\theta_i) + L \leq \theta \leq \max_{i\in\mathcal{J}}(\theta_i) + U \tag{8}$$

can be recast as an ELCP. First of all we introduce a dummy variable $\gamma = \max_{i\in\mathcal{J}} \theta_i$. Then (8) reduces to the linear inequality

$$\gamma + L \leq \theta \leq \gamma + U \;\; , \tag{9}$$

which already fits the ELCP format. Let us now look at the equation $\gamma = \max_{i\in\mathcal{J}} \theta_i$. This can be recast as

$$\gamma \geq \theta_i \quad \text{for all } i \in \mathcal{J} \;\; , \tag{10}$$

where for at least one index $i \in \mathcal{J}$ equality should hold, i.e.

$$\prod_{i\in\mathcal{J}} (\gamma - \theta_i) = 0 \;\; . \tag{11}$$

Clearly, equations (9)–(11) constitute an ELCP.

Thus K_{C^θ} can be treated as an ELCP. First we derive the polyhedron that provides the set of solution for the system of linear (in)equalities given by (6). The solution set of the ELCP is obtained as a union of faces of a polyhedron that satisfy the cross-complementarity condition [DES 95].

Example 4. *Consider the TPN $\langle\mathcal{N}^\theta, M_0^\theta\rangle$ displayed in Fig. 4.*

Since $\mathcal{N}^\theta$ is acyclic, $\mathcal{N}^\theta$ is isomorphic with its unfolding. Consider the configuration C that includes t_1, t_2, t_3.

The following $(max, +)$-system of linear inequalities provides the characteristic system of the configuration C.

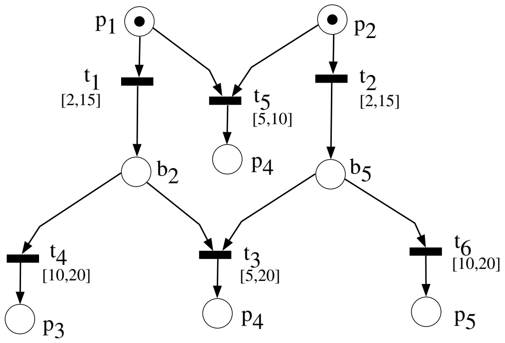

Figure 4. *The TPN of the Example 4.*

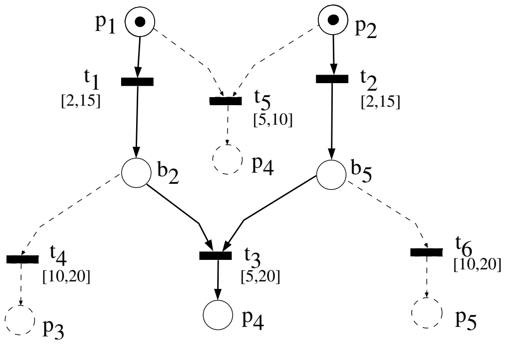

Figure 5. *A configuration of the TPN displayed in Fig. 4.*

$$\begin{cases} 2 \leq \theta_{t_1} \leq 15 \\ 2 \leq \theta_{t_2} \leq 15 \\ \min(\theta_{t_1}, \theta_{t_2}) \leq 10 \\ \max(\theta_{t_1}, \theta_{t_2}) + 5 \leq \theta_{t_3} \leq \max(\theta_{t_1}, \theta_{t_2}) + 20 \\ \theta_{t_3} \leq \theta_{t_1} + 20 \\ \theta_{t_3} \leq \theta_{t_2} + 20 \end{cases} \tag{12}$$

We use the notation x_i for the execution time θ_{t_i} of transition t_i, $i = 1, 2, 3$, $z_1 = \min(x_1, x_2)$, and $y_1 = \max(x_1, x_2)$.

$$A = \left[\begin{array}{cccccc} x_1 & x_2 & x_3 & y_1 & z_1 & \delta \\ & & & & & \\ -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -15 \\ -1 & 0 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 0 & -15 \\ 0 & -1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -10 \\ 0 & 0 & -1 & 1 & 0 & 5 \\ 0 & 0 & 1 & -1 & 0 & -20 \\ -1 & 0 & 1 & 0 & 0 & -20 \\ 0 & -1 & 1 & 0 & 0 & -20 \end{array}\right]$$

We have the characteristic system K_{C^θ} in (12) expressed in the form of an ELCP, (6) and (7).

$$\begin{cases} A \cdot x \geq 0 \\ (A_1 \cdot x)(A_2 \cdot x) = 0 \\ (A_3 \cdot x)(A_4 \cdot x) = 0 \end{cases} \tag{13}$$

where $x = [x_1, x_2, x_3, y_1, z_1, \delta]^T$.

Additionally consider that $x_3 = 23$, which simply means that we want to derive the solution set of K_{C^θ} considering that the observable event t_3 is executed at the time $\theta_{t_3} = 23$.

The solution set of the characteristic system is displayed in Fig. 6 as a union of 2 polytopes (trapezia). The first one has as vertices $(3, 15)$, $(10, 15)$, $(3, 3)$, $(10, 10)$ and the second one has as vertices $(3, 3)$, $(10, 10)$, $(15, 10)$, $(15, 10)$.

6. The method based on constraint propagation

Before formally presenting the second algorithm we introduce first the definition of a time interval configuration.

A time interval configuration $C(\mathbf{I})$ is an untimed configuration $C \in \mathcal{C}$ endowed with time intervals for the execution of the events within the configuration. $\mathbf{I}$ is a vector of dimension $\mid E_C \mid$ that comprises for each event $e \in E_C$ the time interval $I(e)$ in which the event e is assumed executed.

Definition 11. *Given the observation $\mathcal{O}_1^\theta$ and a configuration $C \in \mathcal{C}(\mathcal{O}_1^\theta)$ we have that the time interval configuration $C(\mathbf{I})$ is legal if for any event e_i ($\forall e_i \in E_C$) and*

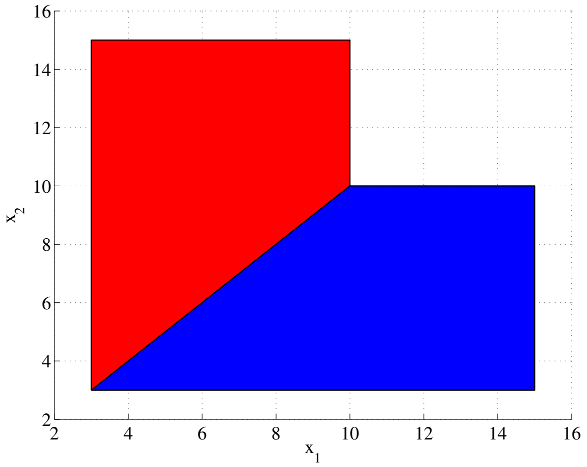

Figure 6. *The projection of the solution set of the characteristic system K_{C^θ} of the Example 4 onto the plane $(\theta_{t_1}, \theta_{t_2})$.*

for any execution time θ_{e_i} of the event e_i ($\forall \theta_{e_i} \in I(e_i)$) there exist execution times for all the other events within the configuration ($\exists \theta_{e_j} \in I(e_j)$ for all $e_j \in E_C \setminus \{e_i\}$) s.t. $\Theta = (\theta_{e_1}, \ldots, \theta_{e_i}, \ldots \theta_{e_{|E_C|}})$ is a solution of the characteristic system K_{C^θ} ($\Theta \in Sol(K_{C^\theta}(\mathcal{O}_1^\theta))$).

Given a hyperbox $\mathbf{I}_\nu \subseteq \mathbf{I}$ denote by $[L_\nu(e), U_\nu(e)]$ the execution time interval for the event e. Then for a conflicting event $\breve{e}$ denote by $L_\nu(\breve{e}) = \max_{e' \in {}^{\bullet\bullet}\breve{e}}(L_\nu(e')) + U^s_{\breve{e}}$ and $U_\nu(\breve{e}) = \max_{e' \in {}^{\bullet\bullet}\breve{e}}(U_\nu(e')) + U^s_{\breve{e}}$ the earliest respectively the latest time when $\breve{e}$ is forced to fire. We have that.

Proposition 3. *$C(\mathbf{I}_\nu)$ is a legal time interval configuration if the following conditions hold true:*

1) $\mathbf{I}_\nu \subseteq \widetilde{\mathbf{I}}$ such that $L_\nu(e) \leq \max_{e' \in {}^{\bullet\bullet}e}(L_\nu(e')) + U^s_e$ and $U_\nu(e) \geq \max_{e' \in {}^{\bullet\bullet}e}(U_\nu(e')) + L^s_e$

2) $\forall \breve{e} \in \breve{E}_C$, $\exists e \in E_C$ s.t. $e \sharp_1 \breve{e}$ and $L_\nu(e) \leq \breve{L}_\nu(\breve{e})$ and $U_\nu(e) \leq \breve{U}_\nu(\breve{e})$

3) $\theta_{obs_1} = \theta_{e^o}$ for $e^o \in E_C$, $\phi(e^o) = l(obs_1)$

4) $\forall e \in {}^\bullet CUT(C) \Rightarrow U_\nu(e) \leq \theta_{obs_1}$

5) $\forall e \in Enbl(C) \Rightarrow \max_{e' \in {}^{\bullet\bullet}e}(L_\nu(e')) + U^s_e \geq \theta_{obs_1}$.

Proof. The proof is as follows. Consider a hyperbox $I_\nu \subseteq \widetilde{I}$ that satisfies the properties $1 - 5$ above.

Denote $\overline{E}_C = E_C \cup \breve{E}_C \cup E^{en}_C$, where E^{en}_C is the set of events that are enabled from $CUT(C)$.

Consider the characteristic system $K_{C^\theta}(\mathcal{O}_1^\theta)$ augmented with the dummy variables $\theta_{\breve{e}}, \theta_e^{en}$ that correspond with the conflicting events and the events that are enabled from $CUT(C)$ where $\theta_{\breve{e}} = \max_{e \in {}^{\bullet\bullet}\breve{e}}(\theta_e) + U_{\breve{e}}^s$ and $\theta_{e^{en}} = \max_{e \in {}^{\bullet\bullet}e^{en}}(\theta_e) + U_{e^{en}}^s$.

Given the execution time $\theta_{\overline{e}_i}$ of $\overline{e}_i \in \overline{E}_C$, denote by $Sol_\nu(\overline{K}_{C^\theta} \mid \theta_{\overline{e}_i})$ the subset of solutions of $\overline{K}_{C^\theta}$ parameterized by $\theta_{\overline{e}_i}$. For $\overline{e}_j \in \overline{E}_C$, denote by $I_\nu(\overline{e}_j \mid \theta_{\overline{e}_i})$ the projection of $Sol_\nu(\overline{K}_{C^\theta} \mid \theta_{\overline{e}_i})$ onto the plane $(\overline{e}_i, \overline{e}_j)$, $I_\nu(\overline{e}_j \mid \theta_{\overline{e}_i}) = [\overline{L}_\nu(\overline{e}_j \mid \theta_{\overline{e}_i}), \overline{U}_\nu(\overline{e}_j \mid \theta_{\overline{e}_i})]$.

Consider $\theta'_{\overline{e}_i} \in I_\nu(\overline{e}_i)$ s.t. $L_\nu(\overline{e}_i) \leq \theta_{\overline{e}_i} \leq \theta'_{\overline{e}_i} \leq U_\nu(\overline{e}_i)$.

We have that:

$$\begin{aligned} 0 \leq \Delta\overline{L}_\nu(\overline{e}_j \mid \theta_{\overline{e}_i}) \leq \Delta\theta_{\overline{e}_i} \\ 0 \leq \Delta\overline{U}_\nu(\overline{e}_j \mid \theta_{\overline{e}_i}) \leq \Delta\theta_{\overline{e}_i} \end{aligned} \tag{14}$$

where $\Delta\overline{L}_\nu(\overline{e}_j \mid \theta_{\overline{e}_i}) = \overline{L}_\nu(\overline{e}_j \mid \theta'_{\overline{e}_i}) - \overline{L}_\nu(\overline{e}_j \mid \theta_{\overline{e}_i})$, $\Delta\overline{U}_\nu(\overline{e}_j \mid \theta_{\overline{e}_i}) = \overline{U}_\nu(\overline{e}_j \mid \theta'_{\overline{e}_i}) - \overline{U}_\nu(\overline{e}_j \mid \theta_{\overline{e}_i})$, and $\Delta\theta_{\overline{e}_i} = \theta'_{\overline{e}_i} - \theta_{\overline{e}_i}$.

By item 2 we have that $\forall \breve{e} \in \breve{E}_C$, $\exists e \in E_C$ s.t. $e \sharp \breve{e}$ and $L_\nu(e) \leq \breve{L}_\nu(\breve{e})$ and $U_\nu(e) \leq \breve{U}_\nu(\breve{e})$.

Claim: $\forall \theta_{\breve{e}} \in I_\nu(\breve{e})$ we have that $L_\nu(e \mid \theta_{\breve{e}}) \leq \theta_{\breve{e}}$.

Proof. (Claim) By item 3 we have that $L_\nu(e) \leq \breve{L}_\nu(\breve{e})$ and $U_\nu(e) \leq \breve{U}_\nu(\breve{e})$ and by (14) $\Delta L_\nu(e \mid \theta_{\breve{e}}) \leq \Delta\theta_{\breve{e}}$. □

Based on this result we have that imposing the constraints due to the conflicting events (45 degree planes) does not modify the projection of the solution set on to the axis. This means that the projection of the solution set $Sol_\nu(K_{C^\theta})$ onto any of the axis is a single interval (and not a union of intervals). This completes the proof. □

In the following we present an algorithm that derives a set of $\mid E_C \mid$-hyperboxes, $\{\mathbf{I}_\nu \mid \nu \in \mathcal{V}\}$ ($\mathcal{V}$ the set of indexes) s.t. for each $\mid E_C \mid$-hyperbox $\mathbf{I}_\nu$, $C(\mathbf{I}_\nu)$ is a legal time interval configuration and the union of the subsets $\{Sol_\nu(K_{C^\theta}) \mid \nu \in \mathcal{V}\}$ that are circumscribed by $\mathbf{I}_\nu$, $\nu \in \mathcal{V}$ is a cover of the entire solution set $Sol(K_{C^\theta})$, i.e.

$$\bigcup_{\nu \in \mathcal{V}} Sol_\nu(K_{C^\theta}) = Sol(K_{C^\theta}), \text{ where } Sol_\nu(K_{C^\theta}) = Sol(K_{C^\theta}) \cap \mathbf{I}_\nu$$

The idea behind developing the algorithm that we propose is as follows. First we calculate the hyperbox $\widetilde{\mathbf{I}}$ that circumscribes $Sol(\widetilde{K}_{C^\theta})$. Then we should impose the timing constraints imposed by the conditions $2-5$ in Proposition 3. We have three kinds of constraints. Denote by $\mathcal{K}_{conf}$, $\mathcal{K}_{obs}^1$, and $\mathcal{K}_{obs}^2$ the set of constraints imposed by the set of conflicting events (condition (2)), the equality constraint required by the observation of the label l_{obs_1} (condition (3)), and respectively the set of constraints that require that the time configuration is complete w.r.t. the time θ_{obs_1} (none of the concurrent parts of the process are left behind in time).

Consider a constraint κ_e on the time interval $\widetilde{I}(e) = [\widetilde{L}(e), \widetilde{U}(e)]$ of an event $e \in E_C$ where:

$$\kappa_e := \Big\{ I'(e) = [L'(e), U'(e)] \mid L'(e) > \widetilde{L}(e) \text{ or } U'(e) < \widetilde{U}(e) \Big\}$$

The set of solutions of $\widetilde{K}_{C^\theta}$ that satisfy κ_e, denoted $Sol(\widetilde{K}_{C^\theta} \wedge \kappa_e)$, is obtained propagating the constraint κ_e forward to its successors and backwards to its predecessors:

– *forward propagation:* for all $e_\upsilon \in e^{\bullet\bullet}$:

$$L'(e_\upsilon) = \max(\widetilde{L}(e) + L^s_{e_\upsilon}, \widetilde{L}(e_\upsilon)) \text{ and } U'(e_\upsilon) = \min(\widetilde{U}(e) + U^s_{e_\upsilon}, \widetilde{U}(e_\upsilon))$$

– *backward propagation:*

i) for all $e_\upsilon \in {}^{\bullet\bullet}e$: $U'(e_\upsilon) = \min(\widetilde{U}(e) - L^s_e, \widetilde{U}(e_\upsilon))$

ii) for each $e_\upsilon \in {}^{\bullet\bullet}e$ s.t. $\widetilde{L}(e) - U^s_e > \widetilde{U}(e_\upsilon)$ consider a different case $\nu \in \mathcal{V}'$:

ii.1) $L'_\nu(e_\upsilon) = \widetilde{L}(e) - U^s_e$

ii.2) for all $e_\iota \in {}^{\bullet\bullet}e$, $e_\iota \neq e_\upsilon : L'_\nu(e_\iota) = \widetilde{L}(e_\iota)$.

The backward propagation of a constraint κ_e may require to split an $\mid E_C \mid$-hyperbox considering different cases. Notice that the number of cases is not bigger than the number of concurrent predecessor events of the event e to whom the constraint κ_e is applied. For each hyperbox $\mathbf{I}_{\nu'}$, $\nu' \in \mathcal{V}'$ the set of constraints is updated since in general it may be that new constraints appear while some of the previous constraints are satisfied. If a constraint cannot be imposed the case is aborted while if the set of constraints is empty the algorithm returns an hyperbox that circumscribes a subset of solutions of K_{C^θ}.

The constraint propagation algorithm works as follows:

1) first step is to impose the constraints of kind $\mathcal{K}^1_{obs}$ and $\mathcal{K}^2_{obs}$ (required by the received observation)

2) the second step is to impose for each $\mid E_C \mid$-hyperbox that results after step 1, the set of constraints $\mathcal{K}_{conf}$. E.g. for $\mathbf{I}_\nu$ consider that $\exists\ \breve{e} \in E_C$ s.t. condition 2 in Proposition 3 is not satisfied. Then for each $e \in E_C$ s.t. $e \sharp_1 \breve{e}$ consider a different case and impose a constraint $\kappa_e := \{L'_{\nu'}(e) = L_{\nu'}(\breve{e})\}$ if $L_{\nu'}(\breve{e}) \leq L_{\nu'}(e)$ or $\kappa_{\breve{e}} = \{U'_{\nu'}(\breve{e}) = U_\nu(e)\}$ if $U_{\nu'}(\breve{e}) \leq U_{\nu'}(e)$.

3) an arbitrary constraint κ_e or $\kappa_{\breve{e}}$ is selected and then it is imposed backwards. If new constraints appear on the time intervals of the predecessor events of e or $\breve{e}$ then one of these constraints is selected and it is imposed further backwards until a decision is achieved. Then constraints are propagated forward for the $\mid E_C \mid$-hyperboxes that are not aborted. The maximum number of different cases that result propagating recursively a constraint backwards is smaller than the size of the maximum set of concurrent events in the configuration

4) a decision is achieved for each case in finite time since the corner points of each $\mid E_C \mid$-hyperbox are rational numbers and each constraint that is applied either reduces one edge of the $\mid E_C \mid$-hyperbox or returns success/abort.

Example 5. *Consider the TPN displayed in Fig. 2 and the part of the unfolding developed up to the discarding time $\hat{\theta} = 17$ (Fig. 3).*

Consider the configuration displayed in Fig. 7 where the dotted part indicates the conflicting events e_1, e_6, e_{12}. Notice that e_1, e_{12} are passive, thus they are not considered.

To deactivate e_6 we have that either e_3 or e_9 happens before the time e_6 is forced to fire. The interval when e_6 is forced to fire is $[3, 7]$. First case is e_3 is executed before e_6 is forced to fire. This condition is satisfied if e_3 is executed in the interval $[4, 7]$ (instead of $[4, 9]$) and e_6 is forced to fire also in $[4, 7]$ (instead of [3,7]). This is achieved if e_5 is executed in the interval $[2, 5]$ or e_8 is executed in the interval $[2, 4]$. The second case requires that e_9 fires before e_6 is forced to fire. Thus e_9 should be executed in the interval $[3, 7]$.

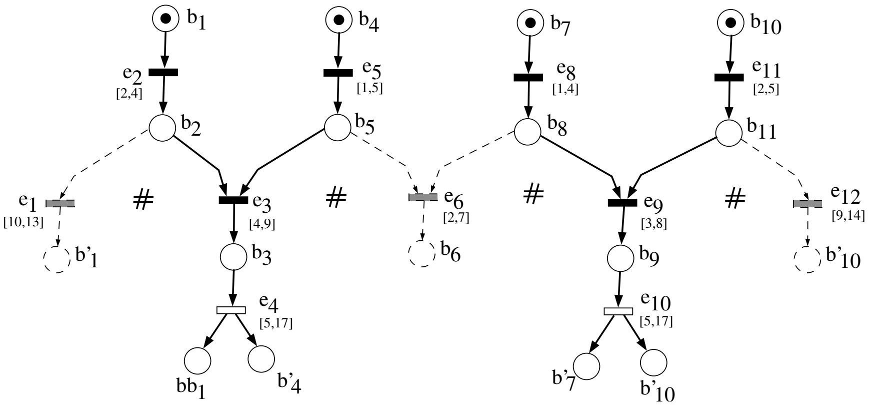

Figure 7. *A configuration of the unfolding displayed in Fig. 3.*

Consider that e_4 is observed at time 13. This constraint is imposed obtaining the following four hyperboxes that provide four legal time interval configurations:

1) $I_1(e_2) = [2, 4]$; $I_1(e_3) = [5, 7]$; $I_1(e_4) = 13$; $I_1(e_5) = [1, 5]$; $I_1(e_8) = [3, 4]$; $I_1(e_9) = [4, 8]$; $I_1(e_{11}) = [2, 5]$;

2) $I_2(e_2) = [2, 4]$; $I_2(e_3) = [5, 7]$; $I_2(e_4) = 13$; $I_2(e_5) = [3, 5]$; $I_2(e_8) = [1, 4]$; $I_2(e_9) = [4, 8]$; $I_2(e_{11}) = [2, 5]$;

3) $I_3(e_2) = [2, 4]$; $I_3(e_3) = [5, 9]$; $I_3(e_4) = 13$; $I_3(e_5) = [1, 5]$; $I_3(e_8) = [2, 4]$; $I_3(e_9) = [4, 7]$; $I_3(e_{11}) = [2, 5]$;

4) $I_4(e_2) = [2, 4]$; $I_4(e_3) = [5, 9]$; $I_4(e_4) = 13$; $I_4(e_5) = [2, 5]$; $I_4(e_8) = [1, 4]$; $I_4(e_9) = [4, 7]$; $I_4(e_{11}) = [2, 5]$;

7. The on-line diagnosis

In the previous sections we have presented the plant diagnosis up to the first observation or in absence of any observation up to the first discarding time. Then the on-line diagnosis is performed calculating the plant behavior up to a new discarding time.

Notice that we can easily extend the time interval configurations since the $\mid E_C \mid$-points Θ_{L_ν} and Θ_{U_ν} that correspond to the lower limits of the execution times $\Theta_{L_\nu} = (L_\nu(e_1) \ldots, L_\nu(e_{|E_C|}))$ respectively to the upper limits of the execution times $\Theta_{U_\nu} = (U_\nu(e_1) \ldots, U_\nu(e_{|E_C|}))$ are solutions of the characteristic system.

Theorem 2. *Given a TPN model $\langle \mathcal{N}^\theta, M_0^\theta \rangle$ we have that:*

1) when an observable event is executed:

$$\mathcal{DR}_{\mathcal{N}^\theta}(\mathcal{O}_n^\theta) = \{\mathtt{F}\} \Leftrightarrow \mathcal{DR}_{\mathcal{N}^\theta}^{po}(\mathcal{O}_n^\theta) = \{\mathtt{F}\}$$

2) for $\hat{\theta}$ the first discarding time after the time when the n^{th} observed event is reported:

$$\mathcal{DR}_{\mathcal{N}^\theta}(\mathcal{O}_{n,\hat{\theta}}^\theta) = \{\mathtt{F}\} \Leftrightarrow \mathcal{DR}_{\mathcal{N}^\theta}^{po}(\mathcal{O}_{n,\hat{\theta}}^\theta) = \{\mathtt{F}\}$$

3) and in absence of any observation, the diagnosis result w.r.t. the detection of the faults that for sure happened calculated any time in between the last observed event and the discarding time is constant, i.e. $\forall \xi \in [\theta_{obs_n}, \hat{\theta})$:

$$\mathcal{DR}_{\mathcal{N}^\theta}(\mathcal{O}_{n,\xi}^\theta) = \{\mathtt{F}\} \Leftrightarrow \mathcal{DR}_{\mathcal{N}^\theta}^{po}(\mathcal{O}_n^\theta) = \{\mathtt{F}\}$$

Proof. The proof is similar to the proof of Theorem 1. □

8. Conclusions

In this paper we have proposed on-line algorithms for the diagnosis of TPN models. We have presented three solutions for implementation. First solution is appropriated for a plant known to have a cyclic behavior and requires to calculate in advance the plant behavior during an operation cycle and then to take the observation generated by the plant into account.

The second solution that we have presented is to update the calculations only at the time a new observation is generated by the plant. This method is suitable for a plant that generates observations with a high frequency.

The third solution that we have proposed is to derive the plant behavior up to a discarding time, i.e. up to a time when in absence of any observation one can discard untimed support traces because they are not consistent with the plant behavior.

The analysis is based on partial orders and it requires to derive for each configuration the solution set of a system of $(\max, +)$-linear inequalities called the characteristic system a configuration.

We have presented two methods for obtaining the solution set of the characteristic system of a configuration: first method derives exactly the entire solution set and is based on the ELCP and the second method is based on constraint propagation and computes a set of hyperboxes such that the subset of solutions of the characteristic system that are circumscribed by the hyperboxes provide a cover of the solution set. Both algorithms are NP-hard problems. Beside the number of events, the number of conflicting events, and the maximum number of predecessors respectively successors of a node in a configuration, the computational complexity of both methods strongly depends on the structure of the system.

However there are a few reasons that allow us to claim that the two methods are computationally more efficient than the ones ([AUR 97], [GHA 05]) presented in the literature. Comparing with the method based on the state class graph computation [GHA 05] our methods have the advantage that not all the interleaving of the concurrent events are considered. Moreover the computational complexity depends in our case on the size of the largest subnet that contains unobservable transitions whereas the computation complexity in [GHA 05] depends on the size of the entire net. The algorithm in [AUR 97] solves a system of $(\max, +)$-inequalities enumerating all the cases for each $\max$-term. This combinatorial approach is known in the literature to be computational less efficient than the ELCP.

Finally notice that for the characteristic system of the configuration considered in Example 5 the ELCP provides 8 subsets while constraint satisfaction only finds 4 subsets. The reason is that each face of a polyhedron that satisfies a cross-complementarity condition provides a legal time interval configuration but the converse is not true. The subset of solutions that is circumscribed by a hyperbox of a legal time interval configuration may be obtained as a union of faces of a polyhedron that satisfy a cross-complementarity condition.

However the set of hyperboxes obtained running the algorithm based on constraint propagation does not allow one to calculate the minimum and maximum time separation between the execution of two events unless a further refinement of the calculations is performed.

We plan to extend the methodology for a distributed setting where the strong assumptions considered in [JIR 06a] to be relaxed.

Acknowledgements

This research was partially supported by the European 6th Framework Network of Excellence "HYbrid CONtrol: Taming Heterogenity and Complexity of Networked Embedded Systems (HYCON)", contract no. FP6-IST-51138 and by the IAP Program initiated by the Belgium State, Prime Minister's Office for Science, Technology and Culture.

G. Jiroveanu was supported by BOF - Ghent University and by a European Union Marie Curie Fellowship during his stay at Delft University of Technology (CTS contract no. HPMT-CT-2001-00278).

9. References

[AUR 97] AURA T., LILIUS J., "Time processes of Time Petri Nets", *18th International Conference on Application and Theory of Petri Nets (ATPN'97) - LNCS*, vol. 1248, 1997, p. 136-155.

[BAR 99] BARONI P., LAMPERTI G., "Diagnosis of large active systems", *Artificial Intelligence*, vol. 101, num. 1, 1999, p. 135-183.

[BEN 03] BENVENSITE A., FABRE E., HAAR S., JARD C., "Diagnosis of asynchronous Discrete Event Systems, a net unfolding approach", *IEEE Transactions on Automatic Control*, vol. 48, num. 5, 2003.

[BER 83] BERTHOMIEU B., MENASCHE M., "An enumerative approach for analyzing Time Petri Nets", *IFIP Congress, Paris*, , 1983.

[CHA 05] CHATAIN T., JARD C., "Time Supervision of Concurrent Systems using Symbolic Unfoldings of Time Petri Nets", *Int. Conference on Formal Modeling and Analysis of Time Systems, Uppsala, Sweden*, , 2005.

[DES 95] DE SCHUTTER B., DE MOOR B., "The Extended Linear Complementarity Problem", *Mathematical Programming*, vol. 71, num. 3, 1995, p. 289-325.

[ESP 94] ESPARZA J., "Model checking using net unfoldings", *Science of Computer Programming*, vol. 23, num. 2, 1994, p. 151-194.

[GHA 05] GHAZEL M., BIGAND M., TOGUYÉNI A., "A temporal-constraint based approach for monitoring of DESs under partial observation", *16th IFAC Triennial World Congress*, Prague, 2005.

[HUL 95] HULGAARD H., BURNS S., "Efficient Timing Analysis of a Class of Petri Nets", *Computer Aided Verification*, , 1995.

[JIR 06a] JIROVEANU G., "Fault diagnosis for large Petri nets", PhD thesis, Ghent University, Gent, Belgium, 2006.

[JIR 06b] JIROVEANU G., BOEL R., DE SCHUTTER B., "Fault Diagnosis for Time Petri Nets", *Workshop on Discrete Event Systems (WODES'O6)*, Ann Arbor, USA, 2006.

[JIR 06c] JIROVEANU G., DE SCHUTTER B., BOEL R., "On-line diagnosis for Time Petri Nets", *International Workshop on Principles of Diagnosis - DX'06*, Spain, 2006.

[MCM 92] MCMILLAN K. L., "Using unfoldings to avoid the state space explosion problem in verification of asynchronous circuits", SPRINGER-VERLAG, Ed., *4th International Workshop on Computer Aid Verification, LNCS*, vol. 663, 1992.

[MER 74] MERLIN P., "A study of recoverability of computer science", PhD thesis, University of California, Irvine, USA, 1974.

[SAM 95] SAMPATH M., SENGUPTA R., LAFORTUNE S., SINNAMOHIDEEN S., TENEKETZIS D., "Diagnosability of Discrete Event Systems", *IEEE Transactions on Automatic*

Control, vol. 40, num. 9, 1995, p. 1555-1575.

[SEM 96] SEMENOV A., YAKOVLEV A., “Verification of asynchronous circuits using Time Petri Net Unfolding”, *Proceedings ACM/IEEE Design Automation Conference*, , 1996.

[YON 98] YONEDA T., RYUBA H., “CTL model checking of Time Petri Nets using geometric regions”, *EICE Transaction on Information & Systems*, vol. E99-D, 1998, p. 1-11.