Delft University of Technology

Delft Center for Systems and Control

Technical report 06-017

Fault diagnosis for time Petri nets*

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

If you want to cite this report, please use the following reference instead:
G. Jiroveanu, R.K. Boel, and B. De Schutter, "Fault diagnosis for time Petri nets," *Proceedings of the 8th International Workshop on Discrete Event Systems (WODES'06)*, Ann Arbor, Michigan, pp. 313–318, July 2006.

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_017.html

Fault Diagnosis for Time Petri Nets

George Jiroveanu*,†
University of Ghent
george.jiroveanu@ugent.be

René K. Boel
University of Ghent
rene.boel@ugent.be

Bart De Schutter
Delft University of Technology
b.deschutter@dcsc.tudelft.nl

***Abstract*— This paper presents an on-line algorithm for fault diagnosis of Time Petri Net (TPN) models. The plant observation is given by a subset of transitions whose occurrence is always reported while the faults are represented by unobservable transitions. The model-based diagnosis uses the TPN model to derive the legal traces that obey the received observation and then checks whether fault events occurred or not. To avoid the consideration of all the interleavings of the unobservable concurrent transitions, the plant analysis is based on partial orders (unfoldings). The legal plant behavior is obtained as a set of configurations. The set of legal traces in the TPN is obtained solving a system of $(max, +)$-linear inequalities called the characteristic system of a configuration. We present two methods to derive the entire set of solutions of a characteristic system, one based on Extended Linear Complementarity Problem and the second one based on constraint propagation that exploits the partial order relation between the events in the configuration.**

I. Introduction

This paper deals with the diagnosis of Discrete Event Systems (DES) where the time is considered as a quantifiable and continuous parameter. Petri Nets (PNs) are considered as model of a DES, and Time Petri Nets model a timed-DES.

In a TPN a transition can be fired after a delay within a given time interval. The execution takes no time to complete. A trace in the plant comprises the transitions that are executed in the TPN model as well as the time of their occurrence.

The plant observation is given by a subset of transitions whose occurrence is always reported and it includes also the time when an observed transition is executed and this is measured with accuracy according to a global clock. The unobservable events are silent, i.e. the execution of an unobservable transition is not acknowledged by the monitoring system. The faults are modeled by a subset of unobservable transitions.

Model-based diagnosis comprises two stages. First the set of traces that are legal from the initial marking and obey the received observation is derived and then the diagnosis result of the plant is obtained checking whether some or all of the legal traces include fault transitions.

Since a transition in a TPN can fire at any time in some interval, TPN models have in general infinite state spaces because a state may have an infinite number of successor states. Methods based on grouping states that are equivalent under a certain equivalence relation into so called *state classes* were proposed in [2], where it was shown that the state class graph of a TPN is finite iff the TPN is bounded. Thus the potentially infinite state space of a TPN can be finitely represented and the analysis of TPN models is computable.

Since the reachability analysis of TPNs based on state-classes becomes computationally infeasible for models of reasonable size (because of the state space explosion due to the interleaving of the unobservable concurrent events) methods based on partial orders were proposed in [1],[3].

The on-line diagnosis algorithm that we propose considers the plant analysis based on time configurations (time processes [1]). 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 time 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 only once s.t. the partial order between the events in the configuration is preserved in the order in which they appear in the trace.

To derive the entire set of all legal time configurations requires to solve a $(\max, +)$-linear system of inequalities called the characteristic system of the configuration.

We present two methods to derive the entire solution set of the characteristic system of a configuration that avoid the explicit consideration of all the cases for each max-term in the characteristic system of $(\max, +)$-linear inequalities (notice that the enumeration of all possible max-elements would imply to interleave concurrent events which is exactly what we wanted to avoid by using partial orders).

The first method uses the Extended Linear Complementarity Problem (ELCP) [9] 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 uses the concept of time interval configuration. A time interval configuration is an untimed configuration endowed with time intervals for the execution of the events within

* Supported by a European Union Marie Curie Fellowship during his stay at Delft University of Technology (CTS contract no. HPMT-CT-2001-0028)

† Currently with TRANSELECTRICA SA, Craiova, Romania, george.jiroveanu@transelectrica.ro

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.

We derive for each untimed configuration a set of hyperboxes of dimension equal with 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 II we introduce the definitions and the notation used in the paper. In Section III the diagnosis of TPNs is formally described while Section IV presents the analysis of TPN based on partial orders. The ELCP is presented in Section V and in Section VI we present the method based on constraint propagation. Section VII concludes the paper with final remarks and further work.

The reader is assumed to be familiar with net unfoldings [7] and the analysis of TPN based on state classes [2].

II. Notation and definitions

A. 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 $Pre(p,t) : \mathcal{P} \times \mathcal{T} \to \{0,1\}$ and $Post(t,p) : \mathcal{T} \times \mathcal{P} \to \{0,1\}$ 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, $M : \mathcal{P} \to \mathbb{N}$, 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 $i = 1, \ldots, k$, $M_{i-1} \geq Pre(\cdot, t_i)$. $M_0 \xrightarrow{\tau} M_k$ denotes that the sequence τ may fire at M_0 yielding M_k.

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 reachable marking M.

B. Occurrence nets

Definition 1: Given a PN $\mathcal{N} = (\mathcal{P}, \mathcal{T}, F)$ the immediate dependence relation $\prec_1 \subset (\mathcal{P} \times \mathcal{T}) \cup (\mathcal{T} \times \mathcal{P})$ is defined as:

$\forall (a,b) \in (\mathcal{P} \times \mathcal{T}) \cup (\mathcal{T} \times \mathcal{P}) : a \prec_1 b$ if $F(a,b) \neq 0$

Define $\prec$ as the transitive closure of $\prec_1$ ($\prec = \prec_1^*$).

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: $\forall (t_1, t_2) \in \mathcal{T} \times \mathcal{T} : t_1 \sharp_1 t_2$ 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 $\forall (a,b) \in (\mathcal{P} \cup \mathcal{T}) \times (\mathcal{P} \cup \mathcal{T})$: $a \sharp b$ if $\exists t_1, t_2$ s.t. $t_1 \sharp_1 t_2$ and $t_1 \preceq a$ and $t_2 \preceq b$. The independence relation $\| \subset (\mathcal{P} \cup \mathcal{T}) \times (\mathcal{P} \cup \mathcal{T})$ is defined as $\forall (a,b) \in (\mathcal{P} \cup \mathcal{T}) \times (\mathcal{P} \cup \mathcal{T}) : a \| b \Rightarrow \neg(a \sharp b) \wedge (a \not\preceq b) \wedge (b \not\preceq a)$.

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: *i)* $\phi(\mathcal{P}) \subseteq \mathcal{P}'$, *ii)* $\phi(\mathcal{T}) \subseteq \mathcal{T}'$, and *iii)* $\forall t \in \mathcal{T}$, the restriction of ϕ to ${}^\bullet t$ is a bijection between ${}^\bullet t$ and ${}^\bullet \phi(t)$ and $\forall t \in \mathcal{T}$, 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 $O = (B, E, \prec_1)$ s.t.: *i)* $\forall a \in B \cup E : \neg(a \prec a)$ (acyclic), *ii)* $\forall a \in B \cup E : \mid \{b : a \prec b\} \mid < \infty$ (well-formed), and *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 $a \prec_1 b \Rightarrow a \in B_C \cup E_C$
iv) $\mathtt{min}_{\preceq}(C) = \mathtt{min}_{\preceq}(O)$.

For a configuration $C \in \mathscr{C}$ denote by $\langle E_C \rangle_{\preceq}$ the set of strings that are linearizations of $(E_C, \preceq)$ where a string $\sigma = e_1 e_2 \ldots e_\upsilon$ 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 $\iota < \lambda$.

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 B of a PN $\langle \mathcal{N}, M_0 \rangle$ is a pair $B = (O, \phi)$ 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 set of all the 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 $\mathscr{B}, \mathscr{B}'$ of PN $\langle \mathcal{N}, M_0 \rangle$ then $\mathscr{B}' \subseteq \mathscr{B}$ if there exists an injective homomorphism $\varphi : \mathscr{B}' \to \mathscr{B}$ s.t. $\varphi(\mathtt{min}(\mathscr{B}')) = \mathtt{min}(\mathscr{B})$ 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 $\mathscr{U}_{\mathcal{N}}(M_0)$ [7]. Denote by $\mathscr{C}$ the set of all the configurations C in $\mathscr{U}_{\mathcal{N}}(M_0)$.

C. 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 $I^s : \mathcal{T} \to \mathscr{I}(\mathbb{Q}^+)$, $I^s(t) = [L_t^s, U_t^s]$, $L_t^s, U_t^s \in \mathbb{Q}^+$, representing the set of all possible time delays associated to transition $t \in \mathcal{T}$.

In a TPN $\langle \mathcal{N}^\theta, M_0^\theta \rangle$ we say that a transition t becomes enabled at the time θ_t^{en} then the clock attached to t is started

and the transition t can and must fire at some time $\theta_t \in [\theta_t^{en} + L_t^s, \theta_t^{en} + U_t^s]$, 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 $\theta_t^{en} + U_t^s$.

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:
   a) if $t'' \neq t$ and $M \geq Pre(\cdot, t'')$ then $FI(t'') = [\max(\theta_{t''}^{en} + L_{t''}^s, \theta_t), \theta_{t''}^{en} + U_{t''}^s]$
   b) else $\theta_{t''}^{en} = \theta_t$ and $FI(t'') = [\theta_{t''}^{en} + L_{t''}^s, \theta_{t''}^{en} + U_{t''}^s]$.

A time trace $\tau^\theta = S_0 \xrightarrow{\langle t_1, \theta_{t_1} \rangle} S_1 \ldots \xrightarrow{\langle t_\upsilon, \theta_{t_\upsilon} \rangle} S_\upsilon$ is legal in a TPN if it satisfies the condition: $\forall \iota = 0, \ldots, \upsilon - 1$, $\exists \theta_{t_{\iota+1}}$ s.t. $S_\iota \xrightarrow{\langle t_{\iota+1}, \theta_{t_{\iota+1}} \rangle} S_{\iota+1}$. 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. $\mathcal{L}_{\mathcal{N}^\theta}(M_0^\theta) = \left\{ \tau \mid \exists \tau^\theta \in \mathcal{L}_{\mathcal{N}^\theta}^\theta(M_0^\theta) \right\}$.

III. 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 $l_o : \mathcal{T} \to \Omega_o \cup \{\varepsilon\}$ where Ω_o is a set of labels and ε is the empty label. $l_o(t) = \varepsilon$ 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, $\mathcal{T}_f \subseteq \mathcal{T}_{uo}$; $l_f : \mathcal{T}_{uo} \to \Omega_f \cup \{\varepsilon\}$ is the fault labeling function (Ω_f is a set of labels and ε is the empty label); $l_f(t) = \varepsilon$ if $t \in \mathcal{T}_{uo} \setminus \mathcal{T}_f$ and $l_f(t) \in \Omega_o$ if $t \in \mathcal{T}_f$
8) the faults are unpredictable, i.e. $\forall t \in \mathcal{T}_f$, $\exists t' \in \mathcal{T} \setminus \mathcal{T}_f$ s.t. *i)* ${}^\bullet t' \subseteq {}^\bullet t$ and *ii)* $L_{t'}^s \leq U_t^s$.

Denote $\mathcal{O}_n^\theta = \langle obs_1, \theta_{obs_1} \rangle, \ldots, \langle obs_n, \theta_{obs_n} \rangle$ the observation of n events executed in the plant, 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.

$\mathcal{L}_{\mathcal{N}^\theta}^\theta(M_0^\theta, \mathcal{O}_n^\theta)$ 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 $\mathcal{O}_n^\theta$.

We say that a time trace τ^θ obeys the observation $\mathcal{O}_n^\theta$ if :

1) the last transition in τ^θ is executed at the time θ_{obs_n}
2) $l_o(\tau) = \mathcal{O}_n$ (the untimed support τ of the legal trace τ^θ obeys the untimed observation support trace $\mathcal{O}_n$)
3) and for $k = 1, \ldots, n$, $\theta_{t_k^o} = \theta_{obs_k}$ with $l_o(t_k^o) = obs_k$ and n the number of observed events in $\mathcal{O}_n^\theta$ (the execution time $\theta_{t_k^o}$ of each observable transition t_k^o in τ^θ is equal with the time θ_{obs_k} that was reported).

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

$$\mathcal{D}_{\mathcal{N}^\theta}(\mathcal{O}_n^\theta) = \left\{ \tau_f \mid \tau_f = l_f(\tau) \wedge \tau^\theta \in \mathcal{L}_{\mathcal{N}^\theta}^\theta(M_0^\theta, \mathcal{O}_n^\theta) \right\} \quad (1)$$

The plant diagnosis indicates that a fault for sure happened if all the traces contain fault events (i.e. $\mathcal{D}_{\mathcal{N}^\theta}(\mathcal{O}_n^\theta)$ does not contain the empty string ε). If $\mathcal{D}_{\mathcal{N}^\theta}(\mathcal{O}_n^\theta)$ contains only the empty string ε then the plant is in a normal state. Otherwise the diagnosis $\mathcal{D}_{\mathcal{N}^\theta}(\mathcal{O}_n^\theta)$ indicates that a fault could have happened but did not necessarily happen. These correspond with the diagnoser states *fault, normal* and respectively *uncertain* [8].

IV. The analysis based on partial orders

The partial order reduction techniques developed for untimed PN [7] are shown in [1],[3] 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, that is for each $e \in E_C$ consider a time value $\theta_e \in \mathbf{T}$ ($\mathbf{T}$ 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 with a valuation $\Theta \in \mathbf{T}^{|E_C|}$ of the execution time for its events is called a time configuration (time process in [1]) of the TPN model.

A time configuration is legal if there is a legal trace $\tau^\theta \in \mathcal{L}_{\mathcal{N}^\theta}^\theta(M_0^\theta)$ in the TPN $\langle \mathcal{N}^\theta, M_0^\theta \rangle$ whose untimed support is a linearization of the partial order relation of the events in the configuration while the execution time θ_t of every transition t considered in the trace τ^θ is identical with the valuation of the event $e \in E_C$ s.t. $\phi(e) = t$.

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_t^s that corresponds in the original TPN to transition t s.t. $\phi(e) = t$.

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

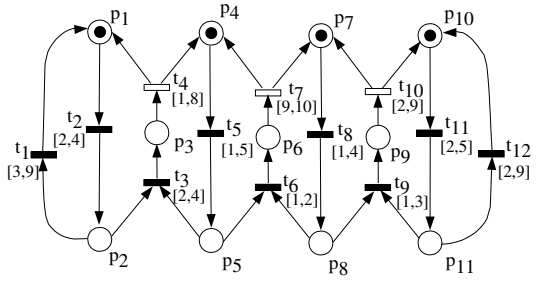

Fig. 1.

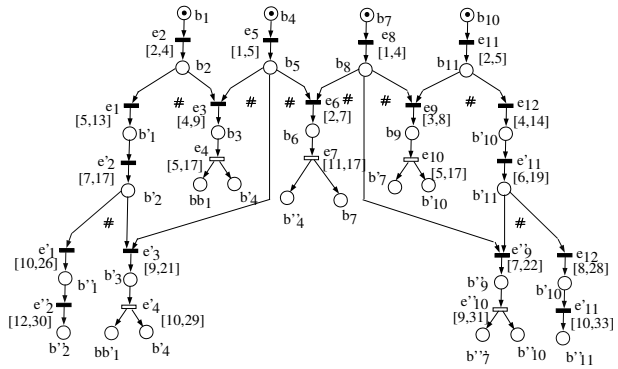

Fig. 2.

- B_C is the set of places, E_C is the set of transitions, and $\preceq$ is the incidence function
- $\mathtt{min}_{\preceq}(\mathscr{U}_{\mathscr{N}})$ is the initial marking (the tokens "arrive" in these places at the time when the process starts)
- $I^s : E_C \to \mathscr{I}(\mathbf{T}_+)$, $I^s(e) = I^s(t)$ with $t = \phi(e)$.

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

$$\widetilde{K}_{C^\theta} = \begin{cases} \max\limits_{e' \in {}^\bullet\bullet e}(\theta_{e'}) + L^s_e \le \theta_e \le \max\limits_{e' \in {}^\bullet\bullet e}(\theta_{e'}) + U^s_e \\ \qquad\qquad\qquad\qquad \text{for } e \in E_C \end{cases} \tag{2}$$

where in (2) ${}^\bullet\bullet e = \emptyset$ implies $\max_{e' \in {}^\bullet\bullet e}(\theta_{e'}) = 0$.

Proposition 1: $\forall \tau^\theta \in \mathscr{L}^\theta_{\mathscr{N}^\theta}(M^\theta_0)$ 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 $\phi(e_i) = t_i$ for $i = 1, \ldots, |\, E_C \,|$.

Proof: The proof is straightforward. □

Denote by $Sol(\widetilde{K}_{C^\theta})$ the set of all solutions of $\widetilde{K}_{C^\theta}$. The $|\, E_C \,|$-hyperbox $\widetilde{I}$ that circumscribes $Sol(\widetilde{K}_{C^\theta})$ is easily obtained as: $\forall e \in E_C$, $\widetilde{I}(e) = [\widetilde{L}(e), \widetilde{U}(e)]$ with $\widetilde{L}(e) = \max_{e' \in {}^\bullet\bullet e}(\widetilde{L}(e')) + L^s_e$ and $\widetilde{U}(e) = \max_{e' \in {}^\bullet\bullet e}(\widetilde{U}(e')) + U^s_e$ where $\forall e \in E_C$ s.t. ${}^\bullet\bullet e = \emptyset$, $\widetilde{L}(e) = L^s_e$ and $U(e) = U^s_e$.

Let the first observation be $\mathscr{O}^\theta_1 = \langle obs_1, \theta_{obs_1} \rangle$. Consider the set of configurations $\mathscr{C}(\mathscr{O}^\theta_1)$ s.t. $C \in \mathscr{C}(\mathscr{O}^\theta_1)$ if:

1) E_C contains only one event e^o s.t. $\phi(e^o) \in \mathscr{T}_o$
2) $\phi(e^o) = t^o$, $\ell(t^o) = obs_1$ and $\theta_{obs_1} \in \widetilde{I}(e^o)$
3) $\forall e \in {}^\bullet CUT(C) \Rightarrow \widetilde{L}(e) \le \theta_{obs_1}$
4) $\forall e \in ENABLED(C) \Rightarrow \widetilde{U}(e) > \theta_{obs_1}$

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

We cannot claim yet that for $C \in \mathscr{C}(\mathscr{O}^\theta_1)$ 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.

Consider the TPN displayed in Fig. 1. Static intervals are attached to each transition. The observable transitions are t_4, t_7 and t_{10} and they emit the same label. Transitions t_3 and t_9 model faults. In Fig. 2 a part of the unfolding $\mathscr{U}_{\mathscr{N}}(M_0)$ is displayed where attached to each event $e \in E$ the interval $\widetilde{I}(e) = [\widetilde{L}(e), \widetilde{U}(e)]$ is displayed.

Denote by $\breve{E}_C$ the set of conflicting events of a configuration $C \in \mathscr{C}$. $\breve{E}_C$ 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 $C^\theta \in \mathscr{C}(\mathscr{O}^\theta_1)$ is obtained adding to $\widetilde{K}_{C^\theta}$ inequalities regarding the conflicting events as well as equalities and inequalities imposed by the received observation (e.g. the observed events are executed at the time given by the received observation and the enabled events have their earliest execution time bigger than the time of the last observation):

$$K_{C^\theta} = \begin{cases} \max\limits_{e' \in {}^\bullet\bullet e}(\theta_{e'}) + L^s_e \le \theta_e \le \max\limits_{e' \in {}^\bullet\bullet e}(\theta_{e'}) + U^s_e \; e \in E_C \\ \min\limits_{e' \sharp_1 \breve{e}}(\theta_{e'}) \le \max\limits_{e'' \in {}^\bullet\bullet \breve{e}}(\theta_{e''}) + U^s_{\breve{e}} \qquad \breve{e} \in \breve{E}_C \\ \theta_{e^o} = \theta_{obs_1} \text{ for } \phi(e^o) = t^o \text{ and } \ell(t^o) = obs_1 \\ \theta_{e'} \ge \theta_{obs_1} \text{ for all } e' \in ENABLED(C) \end{cases} \tag{3}$$

The extension to a sequence of observed events $\mathscr{O}^\theta_n = \langle obs_1, \theta_{obs_1} \rangle, \ldots, \langle obs_n, \theta_{obs_n} \rangle$ is straightforward.

Proposition 2: Given the observation generated by the plant $\mathscr{O}^\theta_n$ we have that $\tau^\theta \in \mathscr{L}^\theta_{\mathscr{N}^\theta}(\mathscr{O}^\theta_n)$ iff:

1) $\tau = \phi(\sigma)$, $\sigma \in \langle E_C \rangle_\preceq$ and $C \in \mathscr{C}(\mathscr{O}^\theta_n)$
2) Θ is a solution of $K_{C^\theta}(\mathscr{O}^\theta_n)$
3) $\forall e \in E_C \setminus \{e^o_n\} \Rightarrow \theta_e \le \theta_{obs_n}$.

Proof: The proof can be found in [6]. □

Thus the plant diagnosis is obtained deriving the set of solutions of the characteristic system $K_{C^\theta}(\mathscr{O}^\theta_n)$ for each configuration $C \in \mathscr{C}(\mathscr{O}^\theta_n)$. A naive approach to derive the entire set of solutions of $K_{C^\theta}(\mathscr{O}^\theta_n)$ is to simply enumerate all cases for each max term, i.e. for $\max(\theta_{e_i}, \theta_{e_j})$ to consider a case when $\theta_{e_i} \le \theta_{e_j}$ and the second case $\theta_{e_j} \le \theta_{e_i}$ and then to simply solve each system of linear inequalities. By doing this we interleave concurrent events which is exactly what we wanted to avoid by using partial orders.

Next we present two methods to derive the solution set of the characteristic system of a configuration in a more efficient way. The first method is based on the ELCP while the second method is based on propagating constraints of the execution time intervals of the events in a configuration.

V. The method based on ELCP

The ELCP is defined as follows [9]:

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

such that

$$Ax \geq c, \qquad Bx = d \tag{4}$$

$$\sum_{j=1}^{m} \prod_{i \in \psi_j} (Ax - c)_i = 0 \ . \tag{5}$$

Condition (5) can be interpreted as follows. Since $Ax \geq c$, all the terms in (5) are nonnegative. Hence,(5) 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 [9] we have developed an algorithm to find *all* solutions of an ELCP. 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

$$\max_{i \in \mathscr{J}}(\theta_i) + L \leq \theta \leq \max_{i \in \mathscr{J}}(\theta_i) + U \tag{6}$$

can be recast as an ELCP. First of all we introduce a dummy variable $\gamma = \max_{i \in \mathscr{J}} \theta_i$. Then (6) reduces to the linear inequality

$$\gamma + L \leq \theta \leq \gamma + U \ , \tag{7}$$

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

$$\gamma \geq \theta_i \quad \text{for all } i \in \mathscr{J} \ , \tag{8}$$

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

$$\prod_{i \in \mathscr{J}} (\gamma - \theta_i) = 0 \ . \tag{9}$$

Clearly, equations (7)–(9) constitute an ELCP.

Thus $K_{C^\theta}(\mathscr{O}_n^\theta)$ can be treated as an ELCP. First we derive the polyhedron that provides the set of solutions for the system of linear (in)equalities given by (4). The solution set of the ELCP is obtained as a union of faces of a polyhedron that satisfy the cross-complementarity condition [9].

VI. The method based on constraint propagation

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

A time interval configuration $C(I)$ is an untimed configuration $C \in \mathscr{C}$ endowed with time intervals for the execution of the events within the configuration. 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 that was executed.

Definition 9: Given the observation $\mathscr{O}_1^\theta$ and a configuration $C \in \mathscr{C}(\mathscr{O}_1^\theta)$ we have that the time interval configuration $C(I)$ is legal if for any event e_i $(\forall e_i \in E_C)$ and for any execution time θ_{e_i} of the event e_i $(\forall \theta_{e_i} \in I(e_i))$ there exists 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 $K_{C^\theta}(\mathscr{O}_1^\theta)$ $(\Theta \in Sol(K_{C^\theta}))$.

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

Proposition 3: $C(I_\nu)$ is a legal time interval configuration if the following conditions hold true:

1) $I_\nu \subseteq \widetilde{I}$ such that $L_\nu(e) \leq \max_{e' \in {}^\bullet\bullet e}(L_\nu(e')) + U_e^s$ and $U_\nu(e) \geq \max_{e' \in {}^\bullet\bullet e}(U_\nu(e')) + L_e^s$
2) $\forall \check{e} \in \check{E}_C$, $\exists e \in E_C$ s.t. $e \sharp_1 \check{e}$ and $L_\nu(e) \leq \check{L}_\nu(\check{e})$ and $U_\nu(e) \leq \check{U}_\nu(\check{e})$.
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) $\forall e \in ENABLED(C) \Rightarrow$ $\max_{e' \in {}^\bullet\bullet e}(L_\nu(e')) + U_e^s \geq \theta_{obs_1}$.

Proof: The proof is lengthy and is omitted. □

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

The idea behind developing the algorithm that we propose is as follows. First we calculate the hyperbox $\widetilde{I}$ 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 $\mathscr{K}_{conf}$, $\mathscr{K}'_{obs}$, and $\mathscr{K}''_{obs}$ 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:
$\kappa_e := \left\{ I'(e) = [L'(e), U'(e)] \mid L'(e) > \widetilde{L}(e) \text{ or } U'(e) < \widetilde{U}(e) \right\}$

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 $e_\upsilon \in e^{\bullet\bullet}$:
  $L'(e_\upsilon) = \max(\widetilde{L}(e) + L_{e_\upsilon}^s, \widetilde{L}(e_\upsilon))$ and
  $U'(e_\upsilon) = \min(\widetilde{U}(e) + U_{e_\upsilon}^s, \widetilde{U}(e_\upsilon))$
- *backward propagation:*
  i) for all $e_\upsilon \in {}^{\bullet\bullet}e$: $U'(e_\upsilon) = \min(\widetilde{U}(e) - L_e^s, \widetilde{U}(e_\upsilon))$
  ii) for each $e_\upsilon \in {}^{\bullet\bullet}e$ s.t. $\widetilde{L}(e) - U_e^s > \widetilde{U}(e_\upsilon)$ consider a different case $\nu \in \mathscr{V}'$:
  ii.1) $L'_\nu(e_\upsilon) = \widetilde{L}(e) - U_e^s$
  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 $I_{\nu'}$, $\nu' \in \mathscr{V}'$ the set of constraints is updated since in general it may be that new constraints appear while some of the previous

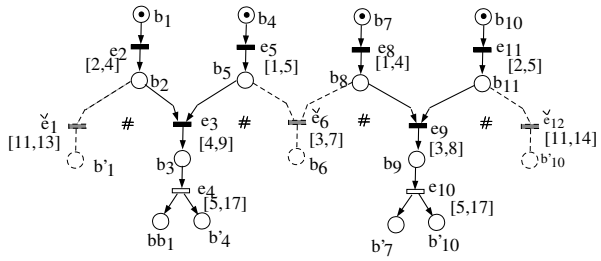

Fig. 3.

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 $\mathscr{K}'_{obs}$ and $\mathscr{K}''_{obs}$ (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 $\mathscr{K}_{conf}$. E.g. for I_v consider that $\exists\ \check{e} \in E_C$ s.t. condition 2 in Proposition 3 is not satisfied. Then for each $e \in E_C$ s.t. $e \sharp_1 \check{e}$ we consider a case and try to impose a constrain $\kappa_e := \left\{L'_{v'}(e) = L_{v'}(\check{e})\right\}$ if $L_{v'}(\check{e}) \leq L_{v'}(e)$ or a constraint $\kappa_{\check{e}} := \left\{U'_{v'}(\check{e}) = U_v(e)\right\}$ if $U_{v'}(\check{e}) \leq U_{v'}(e)$.
3) an arbitrary constraint κ_e or $\kappa_{\check{e}}$ is selected and then it is imposed backwards. If new constrains appear on the time intervals of the predecessor events of e or $\check{e}$ then one of these constraints is selected and it is imposed further backwards until a decision is achieved. Then constrains 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 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 1: Consider for the configuration C displayed in Fig. 3 that the first observation is received at the time 13 and consider the case when e_4 is the event that was observed. Let $\kappa'_{e_4} = \{\theta_{e_4} = 13\}$. κ'_4 is propagated backwards and a new constraint κ'_{e_3} appears where $\kappa''_{e_3} = \{I_{e_4} = [5,9]\}$. κ'_{e_3} is propagated backwards but no new constraints appears. Then e_{10} is required to be executed after $\theta_{e_4} = 13$, i.e. $\kappa''_{e_{10}} = \{\theta_{10} \in [13,17]\}$. $\kappa''_{e_{10}}$ is propagated backwards and a constraint κ_{e_9} appears where $\kappa''_{e_9} = \left\{I_{e_9} = [4,8]\right\}$. κ''_{e_9} is propagated backwards and no new constraint appears.

Then the timing constraints required by the conflicting events $\check{e}_1$ and $\check{e}_{12}$ are satisfied. What is left is the conflicting event $\check{e}_6$. We have that $e_3 \sharp \check{e}_6$ and $e_9 \sharp \check{e}_6$ and $I(e_3) = [5,9]$, $I(e_9) = [4,8]$, and $I(\check{e}_6) = [3,7]$.

We have two cases. First consider $e_3 \sharp \check{e}_6$. We have $\kappa_{\check{e}_6} = \left\{L'_{e_6} = 5\right\}$ and $\kappa_{e_3} = \left\{U'_{e_3} = 7\right\}$. $\kappa_{\check{e}_6}$ is propagated backwards and we have two cases: either $I_1(e_5) = [2,5]$ and $I_1(e_8) = [1,4]$ or $I_2(e_5) = [1,5]$ and $I_2(e_8) = [2,4]$. $\kappa_{e_3} = \left\{U'_{e_3} = 7\right\}$ does not produce new constraints. We obtain two hyperboxes and if we consider the case when $e_9 \sharp \check{e}_6$ we obtain in a similar way another two hyperboxes.

VII. Final remark and perspectives

Both algorithms that we presented are NP-hard problems. Beside the number of events, the number of conflicting events in a configuration, and the maximum number of predecessors resp. 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 ([1], [5]) presented in the literature. Comparing with the method based on the state class graph computation [5] 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 [5] depends on the size of the entire net. The algorithm proposed in [1] 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 above example the ELCP would provide 8 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 the hyperbox of a time interval configuration may be obtained as a union of faces of a polyhedron that satisfy a cross-complementarity condition. This lesser number of subsets provides an advantage in a distributed setting where local agents exchange information for achieving the global consistency of their local diagnosis [6]. The idea is that for each legal time interval configuration $C(I_v)$ the left bottom corner respectively the right top corner of the $\mid E_C \mid$-hyperbox I_v are solutions of the characteristic system the configuration and the constraint propagation algorithm uses only the lower and the upper limits of the execution time intervals as well as the static firing intervals for the events within a configuration.

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 [6] to be relaxed.

References

[1] T. Aura and J. Lilius, Time Processes for Time Petri Nets *ATPN 1997*, LNCS pp. 136-155, Springer Verlag, 1997

[2] B. Berthomieu and M. Menasche, An enumerative approach for analyzing Time Petri Nets *IFIP Congress* Paris, 1983.

[3] T. Chatain and C. Jard, Time Supervision of Concurrent Systems using Symbolic Unfoldings of Time Petri Nets *Int. Conf. on Formal Modeling and Analysis of Time Systems* Uppsala, Sweden, 2005.

[4] E. Fabre, A. Benvensite, S. Haar and C. Jard Distributed monitoring of concurrent and asynchronous systems *JDEDS*, March, 2005
[5] M. Ghazel, M. Bigand and A. Toguyény, A temporal-constraint based approach for the monitoring of DESs under partial observation *Proc. of IFAC Congress*, Prague, 2005
[6] G. Jiroveanu, Fault diagnosis for large Petri Nets *PhD Thesis*, Ghent University, 2006
[7] K. L. McMillan, Symbolic model checking *Kluwer Academic*, 1993
[8] M. Sampath, R. Sengupta, S. Lafortune, S. Sinnamohideen and D. Teneketzis, Diagnosability of Discrete Event Systems *IEEE-T on Automatic Control* Vol. 40(9), 1995, pp. 1555-1575
[9] B. De Schutter and B. De Moor, The Extended Linear Complementarity Problem *Mathematical Programming*, 71(3):289-325, Dec. 1995