Stochastic timed event graphs : bounds, cycle time reachability and marking optimization

Jean-Marie Proth, Nathalie Sauer, Xiaolan Xie

UNITÉ DE RECHERCHE
INRIA-LORRAINE

Rapports de Recherche

1992

ème anniversaire

N° 1763

Programme 5
Traitement du Signal,
Automatique et Productique

STOCHASTIC TIMED EVENT GRAPHS : BOUNDS, CYCLE TIME REACHABILITY AND MARKING OPTIMIZATION

Jean-Marie PROTH
Natalie SAUER
Xiaolan XIE

Septembre 1992

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P. 105
78153 Le Chesnay Cedex
France
Tél.:(1) 39 63 55 11

Graphes d'Evénements Stochastiques : Bornes, Atteignabilité du Temps de Cycle et Optimisation du Marquage

Jean-Marie PROTH, Natalie SAUER, et Xiaolan XIE
Projet SAGEP / INRIA-Lorraine
Technopôle Metz 2000, 4 rue Marconi, 57070 Metz, France

Résumé

Ce papier est consacré à l'évaluation et à l'optimisation des performances des graphes d'événements stochastiques. Les temps de franchissement des transitions sont générés par des variables aléatoires quelconques. Nous proposons d'abord une borne inférieure et une borne supérieure du temps de cycle. Nous montrons qu'un temps de cycle, strictement supérieur au maximum des valeurs moyennes des temps de franchissement, peut être atteint si un nombre suffisant de jetons est disponible dans chaque place. Nous donnons également une condition nécessaire et suffisante pour atteindre un temps de cycle égal au maximum des valeurs moyennes des temps de franchissment. Enfin, nous proposons un algorithme pour résoudre le problème d'optimisation du marquage qui consiste à obtenir un temps de cycle donné tout en minimisant un critère linéaire fonction du marquage initial.

Mots clefs: Graphes d'événements stochastiques, Evaluation des performances, Bornes, P-invariant

STOCHASTIC TIMED EVENT GRAPHS : BOUNDS, CYCLE TIME REACHABILITY AND MARKING OPTIMIZATION

Jean-Marie PROTH, Natalie SAUER, and Xiaolan XIE*

INRIA, Technopôle Metz 2000, 4 rue Marconi, 57070 METZ, FRANCE

ABSTRACT: This paper addresses the performance evaluation and optimization of stochastic timed event graphs. The transitions firing times of such a timed event graph are random variables with general distribution. We first establish an upper bound and a lower bound for the average cycle time of the timed event graph. We prove that any cycle time greater than the greatest mean transition firing time can be reached by putting enough tokens in each place. The necessary and sufficient condition of the reachability of the greatest mean firing time is established. We then address the marking optimization problem which consists in obtaining a given cycle time while minimizing a linear criterion depending on the initial marking.

KEYWORDS: DEDS, Stochastic Timed Petri Nets, Performances, Bounds, Optimization

1. INTRODUCTION

Petri nets have been proven to be an adequate tool for modelling discrete event systems with synchronization, concurrency and common resources. Petri nets have been applied to telecommunication systems, computer systems, manufacturing systems, etc. Excellent surveys can be found in [10, 14].

In this paper we limit ourselves to stochastic timed event graphs which form an elementary class of Petri nets. An event graph is a Petri net in which each place has exactly one input transition and one output transition. A strongly connected event graph has some important properties, specifically: (i) the number of tokens in any elementary circuit is constant, and (ii) the system is deadlock free iff each elementary circuit contains at least one token (see for instance [4, 5, 7, 8]).

In the deterministic case, it has been proven [4, 11] that: (i) the cycle time of an elementary circuit is given by the ratio of the sum of the firing times of the transitions of the circuit by the number of tokens in the circuit; (ii) the cycle time of a strongly connected event graph is equal to the greatest cycle time among the ones of all the elementary circuits. Furthermore, a specified cycle time α being given, algorithms have been proposed in [9] to find initial marking which leads to a cycle time less than α while minimizing a linear criterion.

In the stochastic case, it is no more possible to take advantage of the elementary circuits to evaluate the behaviour of the event graph and to reach a given performance. Previous work mainly focused on ergodicity conditions and performance bounds. Ergodicity conditions

* Address all correspondence to Xiaolan XIE, INRIA, Technopôle Metz 2000, 4 rue Marconi, 57070 Metz, France. E-mail : xie@ilm.loria.fr

have been obtained for timed event graphs [1], for stochastic Petri nets [6] and for max-plus algebra models of stochastic discrete event systems [13]. For a strongly connected timed event graph, it has been proven that an average cycle time exists under some fairly weak conditions (see section 2).

This paper addresses three different issues : the performnce bounds, the cycle time reachability and the marking optimization problem. In section 3, we propose an upper bound and a lower bound for the average cycle time according to the initial marking and compare them with the existing bounds.

Under some weak conditions described in section 2, we prove in section 4 that it is always possible to reach an average cycle time smaller than any given value α which is greater than the greatest average transition firing time. We also establish the necessary and sufficient condition for the reachability of the greatest average transition firing time.

Section 5 addresses the marking optimization problem which consists in finding a marking which minimizes a p-invariant criterion and leads to an average cycle time smaller than a prespecified cycle time. A heuristic algorithm which provides near-optimal solutions is proposed.

2. NOTATIONS AND ASSUMPTIONS

Let $N = (P, T, F)$ be the strongly connected event graph considered. P is the set of places, T is the set of transitions, and $F \subset (P \times T) \cup (T \times P)$ is the set of directed arcs. We denote by M_0 the initial marking of N.

We assume that no transition can be fired by more than one token at any time (i.e. recycled transitions). We further assume that, when a transition fires, the related tokens remain in the input places until the firing process ends. They then disappear and one new token appears in each output place of the transition.

The following notations are used throughout this paper:

$X_t(k) \in \mathbb{R}^+$: time required for the k-th firing of transition t. It is a random variable

$S_t(n)$: instant of the n-th firing initiation of transition t

Γ: set of elementary circuits of N

$M_0(\gamma)$: total number of tokens contained initially in $\gamma \in \Gamma$

$\mu(\gamma) = \sum_{t \in \gamma} X_t(1)$: sum of the firing times of transitions belonging to γ

We assume that the sequences of transitions firing times $\{X_t(k)\}_{k=1}^{\infty}$ for $t \in T$ are mutually independent sequences of independent identically distributed (i.i.d.) integrable random variables.

It was proven in [1] that, under the foregoing assumptions, there exists a positive constant $\pi(M_0)$ such that:

$$\lim_{n\to\infty} S_t(n)/n = \lim_{n\to\infty} E\left[S_t(n)\right]/n = \pi(M_0), \quad \text{a.s.} \ \forall\, t \in T$$

$\pi(M_0)$ is the average cycle time of the event graph.

Since $\{X_t(k)\}_{k=1}^{\infty}$ are sequences of i.i.d. random variables, the index k is often dropped off and we use X_t to denote the firing time of transition t whenever k is not necessary. We further assume that the first and second moments of X_t exists and denote by m_t its mean value and by σ_t its standard deviation, i.e. $m_t = E[X_t]$ and $\sigma_t^2 = E[(X_t - m_t)^2]$.

3. BOUNDS OF THE AVERAGE CYCLE TIME

In this section we assume that the number of tokens as well as their distribution in the strongly connected event graph are known at the initial time. Let M_0 be this marking. We provide a lower bound and an upper bound of the mean cycle time of the system.

We consider an operational mode of the event graph model, called **earliest operational mode (EOM)**, for which transitions fire as soon as they are enabled, provided they are idle. As shown by Chretienne [4], this operation mode leads to the minimal mean cycle time. This mode is used in the following.

3.1. The lower bound

Note that the cycle time which is the solution to the deterministic problem obtained by replacing the random variables which generate the firing times by their mean values is a lower bound of the mean cycle time. Proposition 1 provides a better lower bound of the value of the mean cycle time than the previous one. We denote it by $\underline{\pi}$.

Proposition 1:

The following inequality holds:

$$\pi(M_0) \geq \underset{\gamma\in\Gamma}{\mathrm{Max}}\, E\left[\max\left\{\frac{\mu[\gamma \setminus \{t^*(\gamma)\}] + m_{t^*(\gamma)}}{M_0(\gamma)}, m_{t^*(\gamma)}\right\}\right] = \underline{\pi} \tag{1}$$

where:

$t^*(\gamma)$ is a transition belonging to γ which has the greatest average firing time, i.e.

$$m_{t^*(\gamma)} = \max_{t\in\gamma} m_t$$

$\mu[\gamma \setminus \{t^*(\gamma)\}]$ is the sum of firing times of transitions belonging to γ except $t^*(\gamma)$, i.e.

$$\mu[\gamma \setminus \{t^*(\gamma)\}] = \sum_{t\in\gamma, t\neq t^*(\gamma)} X_t$$

Proof:

Consider an elementary circuit $\gamma = (t_1, p_1, t_2, p_2, \dots t_v, p_v, t_1)$. The following relations hold:

$$S_{t_1}(n) + X_{t_1}(n) \leq S_{t_2}(n + M_0(P_1))$$

....

$$S_{t_v}(n + M_0(P_1) + \dots + M_0(P_{v-1})) + X_{t_v}(n + M_0(P_1) + \dots + M_0(P_{v-1})) \leq S_{t_1}(n + M_0(\gamma))$$

which leads to:

$$S_{t_1}(n+M_0(\gamma))-S_{t_1}(n)\geq\sum_{i=1}^{v}X_{t_i}(n+M_0(P_1)+\ldots+M_0(P_{i-1}))$$

Furthermore, we have:

$$S_{t_1}(n)+X_{t_1}(n)\leq S_{t_1}(n+1)$$

$$\ldots\ldots$$

$$S_{t_1}(n+M_0(\gamma)-1)+X_{t_1}(n+M_0(\gamma)-1)\leq S_{t_1}(n+M_0(\gamma))$$

which yields:

$$S_{t_1}(n+M_0(\gamma))-S_{t_1}(n)\geq\sum_{i=1}^{M_0(\gamma)}X_{t_1}(n+i-1)$$

Combining the above relations:

$$S_{t_1}(n+M_0(\gamma))-S_{t_1}(n)\geq\max\left\{\sum_{i=1}^{v}X_{t_i}(n+M_0(P_1)+\ldots+M_0(P_{i-1})),\ \sum_{i=1}^{M_0(\gamma)}X_{t_1}(n+i-1)\right\}$$

By taking expectation:

$$E\left[S_{t_1}(n+M_0(\gamma))-S_{t_1}(n)\right]$$

$$\geq E\left[X_{t_1}(n)\right]+E\left[\max\left\{\sum_{i=2}^{v}X_{t_i}(n+M_0(P_1)+\ldots+M_0(P_{i-1})),\ \sum_{i=2}^{M_0(\gamma)}X_{t_1}(n+i-1)\right\}\right]$$

Since $X_{t_i}(k)$, for k=1,2,... whatever $t_i \in T$, are i.i.d. random variables, we have:

$$E\left[S_{t_1}(n+M_0(\gamma))-S_{t_1}(n)\right]\geq m_{t_1}+E\left[\max\left\{\sum_{i=2}^{v}X_{t_i},\ \sum_{i=1}^{M_0(\gamma)-1}X_{t_1}(i)\right\}\right]$$

According to Jensen's inequality:

$$E\left[S_{t_1}(n+M_0(\gamma))-S_{t_1}(n)\right]$$

$$\geq m_{t_1}+E\left[\max\left\{\mu[\gamma\setminus\{t_1\}],(M_0(\gamma)-1).m_{t_1}\right\}\right]$$

$$=E\left[\max\left\{\mu[\gamma\setminus\{t_1\}]+m_{t_1},M_0(\gamma).m_{t_1}\right\}\right]$$

By letting $n\to\infty$, we obtain:

$$M_0(\gamma).\pi(M_0)\geq E\left[\max\left\{\mu[\gamma\setminus\{t_1\}]+m_{t_1},M_0(\gamma).m_{t_1}\right\}\right]$$

Thus:

$$\pi(M_0)\geq E\left[\max\left\{\frac{\mu[\gamma\setminus\{t_1\}]+m_{t_1}}{M_0(\gamma)},m_{t_1}\right\}\right]$$

By choosing $t_1 = t^*(\gamma)$:

$$\pi(M_0)\geq E\left[\max\left\{\frac{\mu[\gamma\setminus\{t^*(\gamma)\}]+m_{t^*(\gamma)}}{M_0(\gamma)},m_{t^*(\gamma)}\right\}\right],\quad\forall\gamma$$

which implies that:

$$\pi(M_0) \geq \max_{\gamma\in\Gamma}\left\{E\left[\max\left\{\frac{\mu[\gamma\setminus\{t^*(\gamma)\}] + m_{t^*(\gamma)}}{M_0(\gamma)}, m_{t^*(\gamma)}\right\}\right]\right\}$$

Q.E.D.

3.2. The upper bound

The purpose of this section is to establish an upper bound of the average cycle time of EOM, i.e. π (M_0). It is based on a constrained operating mode in which some transitions are temporally blocked. In the following, we first define this operating model. Properties of this operating mode are then proved and we finally derive an upper bound.

Starting from the initial marking M_0, the constrained operating mode works as follows. First, the initial marking M_0 enables a subset of transitions $T_0 \subset T$. These transitions are fired immediately. Let t_1 be the instant at which the last transitions of T_0 finishes its firing. During the period (0, t_1), other transitions as well as transitions whose firing ends before instant t_1 are blocked. At instant t_1, a new marking M_1 is reached. This marking enables another subset of transitions T_1. As above, these transitions are fired immediately. At instant t_2, all these transition firings end and a new marking M_2 is reached. The process continues. We obtain a sequence of markings {M_i, for i =0, 1, ...} which appear at instant t_i and which enables the firing of a subset of transitions T_i.

The proposition is an important property of this operating mode.

Proposition 2 :

The sequence of markings $\{M_i, \forall\ i \geq 0\}$ as well as that of transition subsets $\{T_i, \forall\ i \geq 0\}$ become K-periodic after finite time which implies that there exist three positive integers i_0, J and K such that

$$M_i = M_{i+J} \text{ and } T_i = T_{i+J} \qquad\qquad \text{for all } i \geq i_0$$

and that each transition is fired exactly K times in any J consecutive periods, i.e. it appears K times in the sequence of subsets $\{T_i, T_{i+1}, ..., T_{i+J-1}\}$ for all $i \geq i_0$.

Proof :

First, notice that the sequence of markings $\{M_i, \forall\ i \geq 0\}$ as well as that of transition subsets $\{T_i, \forall\ i \geq 0\}$ are independent of transition firing times. As a result, these two sequences remain the same in case of deterministic case with all transition firing times equal to 1, i.e. $X_t = 1$ for all $t \in T$.

However, in this deterministic case, the constrained operating mode is exactly the same as the earliest operating mode. M_i is the marking reached at instant i and T_i the subset of transitions T_i initiated at instant i.

Moreover, it was proven in [4] that this earliest operating mode becomes K-periodic after finite time which concludes the proof.

Q.E.D.

Notice that the proof shows that the two sequences $\{M_i, \forall i \geq 0\}$ and $\{T_i, \forall i \geq 0\}$ can be obtained by means of the simulation of a deterministic timed event graph.

These two sequences being obtained, we are now in a position to derive an upper bound. As a matter of fact, the constrained operating mode becomes a renewal process when the K-periodicity of the sequences $\{M_i, \forall i \geq 0\}$ and $\{T_i, \forall i \geq 0\}$ is reached, i.e. $i \geq i_0$. The renewal intervals correspond to $t_{i+J} - t_i$. In each renewal interval, each transition is fired exactly K times. As a result, the average cycle time of the constrained operating mode is equal to $E[t_{i+J} - t_i] / K$ which is an upper bound of $\pi(M_0)$. The exact value of this bound is given by the following proposition.

Proposition 3 :

$$\pi(M_0) \leq \frac{1}{K} \sum_{i=i_0}^{i_0+J-1} E\left[\operatorname*{Max}_{t \in T_i} X_t\right] = \bar{\pi} \tag{2}$$

Proof :

Let $\bar{\pi}$ be the average cycle time of the constrained operating mode. It is obvious that

$\pi(M_0) \leq \bar{\pi}$.

Since $\bar{\pi} = E[t_{i_0+J} - t_{i_0}]/K$ and since

$$t_{i+1} - t_i = \operatorname*{Max}_{t \in T_i} X_t$$

we have:

$$t_{i_0+J} - t_{i_0} = \sum_{i=i_0}^{i_0+J-1} (t_{i+1} - t_i) = \sum_{i=i_0}^{i_0+J-1} \operatorname*{Max}_{t \in T_i} X_t$$

Combining the above equations,

$$\bar{\pi} = \frac{1}{K} E\left[\sum_{i=i_0}^{i_0+J-1} \operatorname*{Max}_{t \in T_i} X_t\right] = \frac{1}{K} \sum_{i=i_0}^{i_0+J-1} E\left[\operatorname*{Max}_{t \in T_i} X_t\right]$$

Q.E.D.

3.3. Comparison with existing bounds

Under the assumption of recycled transitions, it was proven in [2, 3] that:

$$\pi(M_0) \geq \max\left\{\max_{\gamma \in \Gamma} \frac{E[\mu(\gamma)]}{M_0(\gamma)}, \max_{t \in T} m_t\right\} = \underline{\pi}'$$

$$\pi(M_0) \leq \sum_{t \in T} m_t = \bar{\pi}'$$

The following property shows that our bounds are better than $\underline{\pi}'$ and $\bar{\pi}'$..

Property:

$\bar{\pi} \leq \bar{\pi}'$ and $\underline{\pi} \geq \underline{\pi}'$

Proof:

a. From Jensen's inequality

$$
\begin{aligned}
\underline{\pi} &= \max_{\gamma\in\Gamma}\left\{ E\left[\max\left\{ \frac{\mu[\gamma\setminus\{t^*(\gamma)\}]+m_{t^*(\gamma)}}{M_0(\gamma)}, m_{t^*(\gamma)} \right\}\right]\right\} \\
&\geq \max_{\gamma\in\Gamma}\left[\max\left\{ \frac{E\ \mu[\gamma\setminus\{t^*(\gamma)\}]+m_{t^*(\gamma)}}{M_0(\gamma)}, m_{t^*(\gamma)} \right\}\right] \\
&= \max_{\gamma\in\Gamma}\left[\max\left\{ \frac{E[\mu(\gamma)]}{M_0(\gamma)}, m_{t^*(\gamma)} \right\}\right] \\
&= \max\left\{ \max_{\gamma\in\Gamma} \frac{E[\mu(\gamma)]}{M_0(\gamma)}, \max_{\gamma\in\Gamma} m_{t^*(\gamma)} \right\} \\
&= \max\left\{ \max_{\gamma\in\Gamma} \frac{E[\mu(\gamma)]}{M_0(\gamma)}, \max_{t\in T} m_t \right\} \\
&= \underline{\pi}'
\end{aligned}
$$

b. First,

$$
\overline{\pi} = \frac{1}{K}\sum_{i=i_0}^{i_0+J-1} E\left[\operatorname*{Max}_{t\in T_i} X_t\right] \leq \frac{1}{K}\sum_{i=i_0}^{i_0+J-1} E\left[\sum_{t\in T_i} X_t\right] = \frac{1}{K}\sum_{i=i_0}^{i_0+J-1}\left[\sum_{t\in T_i} E[X_t]\right]
$$

From the K-periodicity,

$$
\overline{\pi} \leq \sum_{t\in T} E[X_t] = \overline{\pi}'
$$

Q.E.D.

4. CYCLE TIME REACHABILITY

The purpose of this section is to establish the reachability of a given cycle time when enough tokens are available for each place. More precisely, we show that the cycle time tends to the greatest average firing times C^* (i.e. $C^* = \operatorname*{Max}_{t\in T} m_t$), when the number of tokens tends to infinity.

For this purpose, we first introduce a so-called N-POM operating mode which temporally blocks the firing of some transitions. We then establish bounds of the cycle time when using N-POM. We show that C^* can be approached as closely as we want to when the number of tokens increases. Since the cycle time when using N-POM is obviously greater than the one obtained with the same initial marking when using EOM, the reachability of C^* by using EOM is also established. We also establish necessary and sufficient condition of the reachability of C^*.

4.1. N-POM operation mode

Consider an initial marking which assigns N tokens to each place, i.e. $M_0(p) = N, \forall p \in P$. We define a so-called **N-periodic operation mode**, denoted by **N-POM**, as follows:

(i) Each transition $t \in T$ is fired N times under an earliest operation mode. In other words, the N tokens of each place are used for firing their output transition as soon as possible. Let $\theta_1^N(t)$ the sum of the N first firing times of transition t for $t \in T$. Let $\theta_1^N = \mathop{Max}\limits_{t \in T} \theta_1^N(t)$. Transitions which complete N firings before instant θ_1^N (assuming that firings start at instant 0) are frozen until instant θ_1^N (i.e. further firings are not allowed until instant θ_1^N). As a consequence, the marking of the event graph is also M_0 at instant θ_1^N.

(ii) We restart the same process from instant θ_1^N on. Let θ_2^N the time needed to reach again the marking M_0. The process restarts from instant $\theta_1^N + \theta_2^N$ and the third step takes a time θ_3^N, and so on.

Since some transitions are temporarily frozen when using the N-POM, the related mean cycle time is greater than or equal to the one obtained when using the EOM, assuming that the initial marking is the same in both cases. As a consequence, if the strongly connected event graph reaches a mean cycle time smaller than C when the N-POM is applied, it also reaches a mean cycle time smaller than C when the EOM is applied.

In the following, we focus our attention on the N-POM and derive some properties related to the EOM from this study.

Since $\{X_t(k)\}_{k=1}^{\infty}, \forall t \in T$, are mutually independent sequences of i.i.d. random variables, N-POM is a renewal process. The renewal epochs are $0, \theta_1^N, \theta_1^N + \theta_2^N, \ldots$. During each period, each transition $t \in T$ is fired exactly N times.

From the definition of N-POM, the renewal intervals θ_k^N for k = 1, 2, 3, ... are defined as follows:

$$\theta_k^N = \mathop{Max}\limits_{t \in T} \sum_{i=1}^{N} X_t((k-1)N + i) \tag{3}$$

The following properties are easily derived from the ones of X_t, $t \in T$:

a. The mean value and the standard deviation of the random variables θ_k^N exist,

b. $\theta_1^N, \theta_2^N, \theta_3^N, \ldots$ are mutually independent and identically distributed,

c. $\lim_{L\to+\infty}\left(\sum_{k=1}^{L}\theta_K^N / L\right) = \lim_{L\to+\infty}\left(E\left[\sum_{k=1}^{L}\theta_K^N / L\right]\right) = E\left[\theta_1^N\right]$ with probability 1.

A consequence of these properties is that the mean cycle time of the system exists when N-POM is applied. This cycle time is denoted by C_N and:

$$C_N = E\left[\theta_1^N\right] / N \tag{4}$$

4.2. Bounds of the average cycle time when applying N-POM

In this section, we propose bounds of the average cycle time of the N-POM, i.e. C_N. These bounds show that C_N can be as close as we want it to be from C^* if N is large enough, but finite.

Proposition 3:

Under the previous hypotheses:

$$C^* \le C_N \le C^* + \left[2 / N^{1/3}\right]\sum_{t\in T}\sigma_t \tag{5}$$

The standard deviations σ_t being finite for any $t \in T$, the right hand side of (5) tends to C^* as N tends to infinity. For all $C > C^*$, it is possible to find N so that the mean cycle time of the stochastic timed event graph is less than C. The value N can be obtained by solving the equation $(C^* - C) + [2 / y^{1/3}] \sum_{t\in T} \sigma_t = 0$ which leads to $y^* = [\, 2 \sum_{t\in T} \sigma_t / (C^* - C_0)\,]^3$ and $N = \lceil y^* \rceil$ where $\lceil\bullet\rceil$ denotes the smallest integer greater than or equal to y^*.

Proof:

a. We first prove that $C^* \le C_N$.

Since $\{X_t(k), \forall k\}$ for $t \in T$ are mutually independent sequences of i.i.d., from Jensen's inequality:

$$E\left[\theta_1^N\right] = E\left[\mathcal{Max}_{t\in T}\sum_{i=1}^{N} X_t(i)\right] \ge \mathcal{Max}_{t\in T} E\left[\sum_{i=1}^{N} X_t(i)\right] = N.\mathcal{Max}_{t\in T}\, m_t$$

which yields that :

$$C_N = E\left[\theta_1^N\right] / N \ge \mathcal{Max}_{t\in T}\, m_t = C^*$$

b. We now prove the following inequality:

$$C_N \le C^* + \sum_{t\in T} E\left[\left|\sum_{i=1}^{N}(X_t(i) - m_t)\right| / N\right]$$

From the definition of θ_1^N and C*,

$$E\left[\theta_1^N\right] \le E\left[\mathop{Max}_{t\in T} \sum_{i=1}^{N} (X_t(i) - m_t + C^*)\right] = N.C^* + E\left[\mathop{Max}_{t\in T} \sum_{i=1}^{N} (X_t(i) - m_t)\right]$$

which leads to:

$$E\left[\theta_1^N\right] \le N.C^* + \sum_{t\in T} E\left[\left|\sum_{i=1}^{N} (X_t(i) - m_t)\right|\right]$$

and:

$$C_N = E\left[\theta_1^N\right]/N \le C^* + \sum_{t\in T} E\left[\left|\sum_{i=1}^{N} (X_t(i) - m_t)\right|/N\right] \tag{6}$$

which completes the proof of part b.

c. In the following, we prove that:

$$C_N \le C^* + \left[2/N^{1/3}\right]\sum_{t\in T}\sigma_t$$

Let us set:

$$w_t = \left|\sum_{i=1}^{N} (X_t(i) - m_t)\right| / (N.\sigma_t)$$

Relation (6) can be rewritten as :

$$C_N \le C^* + \sum_{t\in T}\sigma_t E[w_t] \tag{7}$$

Since the random variables $X_t(1)$, $X_t(2)$, ..., $X_t(N)$ are i.i.d., we have:

$$E\left[(w_t)^2\right] = \sum_{i=1}^{N} E\left[(X_t(i) - m_t)^2\right]/(N.\sigma_t)^2 = 1/N \tag{8}$$

Furthermore, for any $\varepsilon \in [0, 1]$:

$$E[w_t] = \int_0^{\varepsilon-} w_t P(dw_t) + \int_{\varepsilon-}^{1-} w_t P(dw_t) + \int_{1-}^{\infty} w_t P(dw_t) \tag{9}$$

But:

$$\int_0^{\varepsilon-} w_t P(dw_t) \le \varepsilon P(0 \le w_t < \varepsilon) \tag{10}$$

$$\int_{\varepsilon-}^{1-} w_t P(dw_t) \le P(\varepsilon \le w_t < 1) \le P(w_t \ge \varepsilon) \tag{11}$$

$$\int_{1-}^{\infty} w_t P(dw_t) \le \int_0^{\infty} (w_t)^2 P(dw_t) = E\left[(w_t)^2\right] \tag{12}$$

Finally, taking into account inequalities (10), (11) and (12), equality (9) leads to:

$$E[w_t] \le \varepsilon P(0 \le w_t < \varepsilon) + P(w_t \ge \varepsilon) + E[(w_t)^2]$$
$$= \varepsilon - \varepsilon P(w_t \ge \varepsilon) + P(w_t \ge \varepsilon) + E[(w_t)^2]$$

$$= \varepsilon + (1 - \varepsilon)\, P\, (w_t \geq \varepsilon) + E[(w_t)^2]$$

According to Chebyshev's inequality, we obtain:

$$E\,[w_t] \leq \varepsilon + (1 - \varepsilon)\, E[(w_t)^2] \,/\, \varepsilon^2 + E[(w_t)^2]$$

Using the result (8):

$$E\,[w_t] \leq \varepsilon + (1 - \varepsilon) \,/\, (N\, \varepsilon^2) + 1 \,/\, N$$

$$\leq \varepsilon + 1 \,/\, (N\, \varepsilon^2)$$

Setting $\varepsilon = 1 / N^{1/3}$, we obtain:

$$E\,[w_t] \leq 2 \,/\, N^{1/3}$$

Combining with relation (7),

$$C_N \leq C_0 + [\, 2 \,/\, N^{1/3}\,] \sum_{t \in T} \sigma_t$$

Q.E.D.

4.3. Reachability of C*

The previous results show that any cycle time $C > C^*$ can be reached by putting enough tokens in each place. The remainder of this section is devoted to the reachability condition of the minimal cycle time C^*.

Proposition 4 :

C^* is reachable iff there exists $t^* \in T$ such that

$$P\left[X_{t^*} = \max_{t \in T} X_t\right] = 1 \tag{13}$$

Furthermore,

(a) If this condition holds, $C_N = C^*$, $\forall\, N \geq 1$

(b) If it does not hold, $\pi\,(M_0) > C^*$ for all M_0

As can be noticed, this proposition claims that C^* is reachable iff there exists a transition whose firing time is always the greatest one.

Proof:

(i) Assume that the condition (13) holds. It follows that

$$C^* = E[X_{t^*}]$$

From the definition of C_N,

$$E\left[C_1\right] = E\left[\max_{t \in T} X_t\right] = E\left[X_{t^*}\right] = C^*$$

Since $C_N \leq C_1$ for all $N > 1$, the property (a) is proven.

(ii) Assume that the condition (13) does not hold. We prove in the following that for any initial marking M_0, the cycle time $\pi\,(M_0) > C^*$.

Let t1 be the transition with the greatest average transition firing time, i.e. $m_{t1} = \underset{t \in T}{\mathrm{Max}}\, m_t$. Since condition (13) does not hold, there exists $t2 \in T$ such that

$$P[X_{t2} > X_{t1}] > 0$$

which implies that there exist $\Delta > 0$ and $\varepsilon > 0$ such that

$$P[X_{t2} \geq X_{t1} + \Delta] = \varepsilon \tag{14}$$

Since the event graph is strongly connected, there exist a directed path from t1 to t2 and another path from t2 to t1. Let N1 (N2) be the total number of tokens initially contained the path from t1 to t2 (from t2 to t1). As in the proof of Proposition 1, it can be shown that

$$S_{t1}(n) \geq S_{t2}(n - N2) + X_{t2}(n - N2) \tag{15}$$

$$S_{t2}(n) \geq S_{t1}(n - N1) + X_{t1}(n - N1) \tag{16}$$

$$S_t(n+k) - S_t(n) \geq \sum_{i=n}^{n+k-1} X_t(i) \quad \forall t \in T, n > 0, k > 0 \tag{17}$$

Without loss of generality, assume that N1 = 0 and N2 = N. Let us consider k consecutive firings of transitions t1. From relation (17),

$$S_{t1}(n+k) - S_{t1}(n) \geq \sum_{i=n}^{n+k-1} X_{t1}(i) \tag{18}$$

From relation (15),

$$S_{t1}(n+k) \geq S_{t2}(n+k-N) + X_{t2}(n+k-N)$$

Combining with relation (16), we obtain

$$S_{t1}(n+k) - S_{t1}(n) \geq X_{t1}(n) + X_{t2}(n+k-N) + S_{t2}(n+k-N) - S_{t2}(n)$$

Combining again with relation (17) for t = t2,

$$S_{t1}(n+k) - S_{t1}(n) \geq X_{t1}(n) + \sum_{i=n}^{n+k-N} X_{t2}(i) \tag{19}$$

From relations (18) and (19),

$$S_{t1}(n+k) - S_{t1}(n) \geq X_{t1}(n) + \mathrm{Max}\left\{ \sum_{i=n+1}^{n+k-1} X_{t1}(i), \sum_{i=n}^{n+k-N} X_{t2}(i) \right\}$$

and :

$$S_{t1}(n+k) - S_{t1}(n) \geq \sum_{i=n}^{n+k-1} X_{t1}(i) + \mathrm{Max}\left\{0, \sum_{i=n}^{n+k-N} (X_{t2}(i) - X_{t1}(i+1)) - \sum_{i=n+k-N+2}^{n+k-1} X_{t1}(i) \right\}$$

By taking expectation,

$$E[S_{t1}(n+k) - S_{t1}(n)] \geq k m_{t1} + E\left[\mathrm{Max}\left\{0, \sum_{i=n}^{n+k-N} (X_{t2}(i) - X_{t1}(i+1)) - \sum_{i=n+k-N+2}^{n+k-1} X_{t1}(i) \right\}\right]$$

By Jensen's inequality,

$$E[S_{t1}(n+k) - S_{t1}(n)] \geq k m_{t1} + E\left[\mathrm{Max}\left\{0, \sum_{i=n}^{n+k-N} (X_{t2}(i) - X_{t1}(i+1)) - (N-2) m_{t1} \right\}\right]$$

Since $m_{t1} = C^*$ and since the random variables $X_{t1}(i)$ for all i are i.i.d.,

$$E[S_{t1}(n+k)-S_{t1}(n)] \geq kC^* + E\left[\mathrm{Max}\left\{0, \sum_{i=n}^{n+k-N}(X_{t2}(i)-X_{t1}(i))-(N-2)C^*\right\}\right] \tag{20}$$

Let us define the following event :

$$A = \{X_{t2}(i)-X_{t1}(i) \geq \Delta \quad \forall n \leq i \leq n+k-N\}$$

From relation (14),

$$P(A) = \varepsilon^{(k-N+1)} \tag{21}$$

From relation (20),

$$E[S_{t1}(n+k)-S_{t1}(n)] \geq kC^* + E\left[\sum_{i=n}^{n+k-N}(X_{t2}(i)-X_{t1}(i))-(N-2)C^* \Big/ A\right]P(A)$$

From the definition of A and relation (21),

$$E[S_{t1}(n+k)-S_{t1}(n)] \geq kC^* + \left((k-N+1)\Delta-(N-2)C^*\right)\varepsilon^{(k-N+1)}$$

Setting $k = (N-1).(1+ C^*/\Delta)$, we obtain

$$E[S_{t1}(n+k)-S_{t1}(n)] \geq kC^* + C^*\varepsilon^{(N-1)C^*/\Delta}$$

Finally, by letting $n \to \infty$, we obtain

$$k\pi(M_0) \geq kC^* + C^*\varepsilon^{(N-1)C^*/\Delta}$$

which implies that $\pi(M_0) > C^*$.

Q.E.D.

5. MARKING OPTIMIZATION

In section 4, we proved that it is always possible to reach a mean cycle time smaller than C with a finite number of tokens, provided $C > C^*$. A prespecified cycle time $C > C^*$ being given, the marking optimization aims at finding an initial marking M_0 which minimizes the value of the p-invariant criterion and leads to an average cycle time less than C.

In the following, we first present a heuristic algorithm for solving the marking optimization problem. Subsection 5.2. is devoted to the evaluation of isolated mean cycle times of all elementary circuits which are needed in applying the heuristic algorithm and subsection 5.3. is a numerical example.

5.1. A heuristic solution to the marking optimization problem

The heuristic algorithm presented hereafter leads to a near-optimal solution to the problem. Its first phase consists in computing the optimal solution to the deterministic problem obtained by assigning to each transition the mean value of the related random variable. We use the algorithm presented in [9] to solve this problem. The second phase of the algorithm is a step-by-step process. At each step of the process, we first evaluate the mean isolated cycle time (or mean cycle time for short) of each elementary circuit $\pi(M_0, \gamma)$

for all $\gamma \in \Gamma$. We select the elementary circuit having the greatest mean cycle time. $\mathcal{P}$ being the set of places of this elementary circuit having the smallest coefficients in the p-invariant, we select as many elementary circuits as possible having a great mean cycle time and whose intersection with $\mathcal{P}$ is not empty, and we add one token in a place belonging to this intersection. The process stops when the average cycle time of the strongly connected event graph is less than C.

ALGORITHM

First phase: Computation of an initial solution

1. Compute the optimal solution M_0 to the problem when using the mean value of the related random variable as the firing time of each transition.

The optimal solution is the one which leads to C for the deterministic problem. We use the algorithms presented in [9] to solve this problem. The optimal solution M_0 is specified as the number of tokens in each place at the initial state.

2. Using the initial random variables to generate the firing times, simulate the system in order to obtain the mean cycle time $\pi(M_0)$.

3. If $\pi(M_0) < C$, stop the computation.

Second phase: Increase adequately the set of tokens

4. Let $\gamma_0 \in \Gamma$ be the elementary circuit having the greatest mean cycle time and $\mathcal{P}$ the set of places belonging to γ_0 and having the smallest coefficients in the p-invariant, $\gamma_1 \in \Gamma - \{\gamma_0\}$ the elementary circuit having the greatest mean cycle time and at least one place in common with $\mathcal{P}$, $\gamma_2 \in \Gamma - \{\gamma_0 \cup \gamma_1\}$ the elementary circuit having the greatest cycle time and at least one place in common with $\mathcal{P} \cap \gamma_1$, and so on until we reach γ_q such that $\mathcal{P} \cap \gamma_1 \cap \gamma_2 \cap ... \cap \gamma_{q-1} \cap \gamma_q = \emptyset$.

The goal of this step is to define one or more places belonging to as many elementary circuits having a great cycle time as possible and having a small coefficient in the criterion.

5. Add one token to a place P belonging to $\mathcal{P} \cap \gamma_1 \cap \gamma_2 \cap ... \cap \gamma_{q-1}$. Let M_0 be the new solution.

6. Simulate the system in order to obtain the mean cycle time $\pi(M_0)$ related to M_0.

7. If $\pi(M_0) \leq C$, M_0 is the near-optimal (or optimal) solution ; otherwise, go to 4.

5.2. Evaluation of mean cycle times of elementary circuits

The mean cycle times of the elementary circuits $\pi(M_0, \gamma)$ have been used in the previous heuristic algorithm. Of course, they can be obtained by simulation. However, as the number of elementary circuits is usually very large, the simulation becomes computationally burdensome. In the following, we derive approximate evaluation of these cycle times from the simulation of the timed event graph as described in step 6 of the heuristic algorithm.

To this end, we first derive an exactly expression (i.e. equation (23)) of the average cycle time of the timed event graph and then derive an approximate expression for the mean isolated cycle time of an elementary circuit (i.e. equation (24)).

Let us first consider an elementary circuit γ over the period [0, r]. We use the following notations :

$W_p(r)$: accumulated waiting time of tokens in place p during [0, r]

$s_t(r)$: accumulated service time of transition t during [0, r]

$n_t(r)$: total firing initiation number of transition t during [0, r]

$n(r)$: maximal firing initiation number over the ones of all transitions in γ

Since a token is either waiting in a place or is being used for the firing of a transition, the following relation holds :

$$\sum_{p\in\gamma} W_p(r) + \sum_{t\in\gamma} s_t(r) = M(\gamma)\, r \tag{22}$$

Consider the following steady state performance index :

$$w_p = \lim_{r\to\infty} W_p(r) / n_{p^\circ}(r)$$

where p° is the output transition of place p. w_p is the average waiting time of tokens in p between two firing of transition p°.

Since the event graph is strongly connected,

$$\lim_{r\to\infty} n_t(r)/n(r) = 1$$

From the ergodicity,

$$\lim_{r\to\infty} r/n(r) = \pi(M_0) \quad \text{and} \quad \lim_{r\to\infty} s_t(r) / n_t(r) = m_t$$

Relation (22) can be rewritten as follows :

$$\sum_{p\in\gamma} \frac{W_p(r)}{n_{p^\circ}(r)} \frac{n_{p^\circ}(r)}{n(r)} + \sum_{t\in\gamma} \frac{s_t(r)}{n_t(r)} \frac{n_t(r)}{n(r)} = M(\gamma) \frac{r}{n(r)}$$

By letting $r \to \infty$, we obtain:

$$\sum_{p\in\gamma} w_p + \sum_{t\in\gamma} m_t = M(\gamma)\, \pi(M_0)$$

thus:

$$\pi(M_0) = \frac{\sum_{p\in\gamma} w_p + \sum_{t\in\gamma} m_t}{M(\gamma)} \tag{23}$$

In general, the RHS term of relation (23) is greater than the mean isolated cycle time of γ since it takes into account the the waiting time of tokens in γ for tokens arrived from outside in case of synchronization. The ideal of approximation is to remove this waiting time.

For this purpose, we evaluate the accumulated waiting time of tokens in place p during the firing of its output transition in [0, r]. We denote by $V_p(r)$ this waiting time. The following steady state performance index is then defined:

$$v_p = \lim_{r \to \infty} Vp(r) \,/\, n_{p^\circ}(r)$$

Finally, the approximation is derived from equation (23) by replacing w_p by v_p:

$$\pi(M_0,\gamma) \approx \frac{\sum_{p\in\gamma} v_p + \sum_{t\in\gamma} m_t}{M(\gamma)} \tag{24}$$

5.3. A numerical example

We illustrate the heuristic algorithm by the following example. The strongly connected event graph is presented in figure 1.

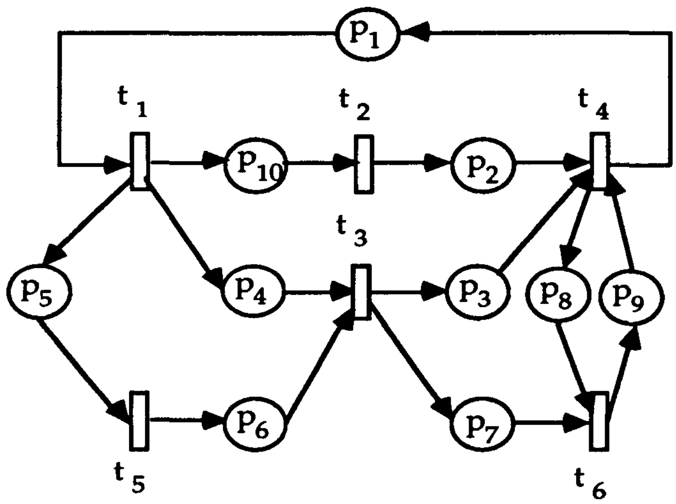

Fig. 1: A strongly connected event graph

The random variables X_1, X_2, X_3, X_4, X_5, X_6 are assigned to the transitions t_1, t_2, t_3, t_4, t_5 and t_6 respectively. Their distributions are the following:

$$X_1: f_1(x) = \begin{cases} 1/10 \text{ if } x \in [0, 10] \\ 0 \text{ otherwise} \end{cases}$$

$$X_2: f_2(x) = \begin{cases} (1 / \beta^{\alpha}\,\Gamma(\alpha)) \,.\, x^{\alpha-1} \exp(-x/\beta) \text{ if } x \geq 0 \\ 0 \text{ otherwise} \end{cases}$$

$\alpha = 2$ and $\beta = 5$

X_3: $f_3(x)$ is the same as f_2, but with $\alpha = 3$ and $\beta = 1$

$X_4 = 5$ (Constant)

$$X_5: f_5(x) = \begin{cases} 10 \exp(-10\,x) \text{ if } x \geq 0 \\ 0 \text{ otherwise} \end{cases}$$

X_6: $P\{X_6 = 3\} = 1/2$ and $P(X_6 = 1) = 1/2$

We choose to minimize the following p-invariant criterion:

$$f(M_0) = 3\,x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + 2\,x_9 + x_{10}$$

where $x_i = M_0(p_i)$ for i =1, 2, ..., 10.

We can see that $C^* = 10$ and we choose $C = 10.1$. The optimal solution of the deterministic problem consists in putting one token in each of the places p_2, p_3, p_4, p_5, p_9, and p_{10}. For this solution, the mean cycle time of the stochastic problem is 12.39 and the p-invariant is equal to 7.

The following steps are given in table 1.

Table 1: Steps of the second stage

Steps	Put one more token in	Mean cycle time	Value of the p-invariant
1	p_{10}	10.76	8
2	p_7	10.63	9
3	p_8	10.52	10
4	p_2	10.12	11
5	p_3	10.09	12

Finally the optimal marking M_0 which leads to an average cycle time less than 10.1 and to a value of the p-invariant which is as small as possible is $M_0 = (0, 2, 2, 1, 1, 0, 1, 1, 1, 2)^T$.

6. CONCLUSION

Three different issues: the performnce bounds, the cycle time reachability and the marking optimization, have been addressed. We first propose an upper bound and a lower bound for the average cycle time which have been proven tighter than the existing bounds.

The most important result is that it is always possible to reach a mean cycle time as close as possible to the greatest mean firing time using a finite marking, assuming that a transition cannot be fired by more that one token at any time. This result holds for any distribution of the transition firing times. We also establish the necessary and sufficient condition for the reachability of the greatest mean firing time.

An efficient heuristic algorithm has been proposed to reach a given cycle time at a low cost (i.e. with a low value of the p-invariant criterion).

BIBLIOGRAPHY

[1] F. BACCELLI, "Ergodic Theory of Stochastic Petri Networks", INRIA Research Report No. 1037, May 1989; also to appear in *Annals of Probability*, 1992.

[2] F. BACCELLI and Z. LIU, "Comparison Properties of Stochastic Decision Free Petri Nets", to appear in *IEEE Transactions on Automatic Control*, 1992.

[3] J. CAMPOS, G. CHIOLA and M. SILVA, "Properties and Performance Bounds for Closed Free Choice Synchronized Monoclass Queuing Networks", *IEEE Transactions on Automatic Control*, Vol. 36, No. 12, pp. 1368-1382, December 1991.

[4] P. CHRETIENNE, "Les réseaux de Petri temporisés", Université Paris VI, Paris, France, Thése d'Etat, 1983.

[5] F. COMMONER, A. HOLT, S. EVEN and A. PNUELI: "Marked Directed Graphs", *Journal of Computer and System Science*, Vol. 5, No. 5, 1971.

[6] P.J. HAAS and G.S. SHEDLER: "Stochastic Petri Nets: Modeling Power and Limit Theorem", *Probability in Engineering and Informational Sciences*, Vol. 5, pp. 477-498, 1991.

[7] H.P. HILLION and J.M. PROTH: "Performance Evaluation of Job-Shop Systems Using Timed Event-Graphs", *IEEE Transactions on Automatic Control*, Vol. 34, No. 1, January 1989.

[8] H.P. HILLION and J.M. PROTH: "Analyse de fabrications non linéaires et répétitives à l'aide des Graphes d'événements Temporisés", *RAIRO*, Vol. 22, No. 2, September 1988.

[9] S. LAFTIT, J.M. PROTH and X.L. XIE: "Optimization of Invariant Criteria for Event Graphs", *IEEE Transactions on Automatic Control*, Vol. 37, No. 5, pp. 547-555, May 1992.

[10] T. Murata, "Petri Nets: Properties, Analysis and Applications," Proceedings of the IEEE, vol. 77, N° 4, April 1989

[11] C.V. RAMAMOORTHY and G.S. HO: "Performance Evaluation of Asynchronous Concurrent Systems using Petri Nets", *IEEE Trans. Software Eng.*, Vol. SE-6, No. 5, pp. 440-449, 1980.

[12] C. RAMCHANDANI: "Analysis of Asynchronous Concurrent Systems by Timed Petri Nets", Lab. Comput. Sci., Mass. Inst. Technol., Cambridge, MA, Tech. Rep. 120, 1974.

[13] J.A.C. RESING, R.E. de VRIES, G. HOOGHIEMSTRA, M.S. KEANE and G.J. OLSDER: "Asymptotic Behavior of Random Discrete Event Systems", *Stochastic Processes and their Applications*, Vol. 36, pp. 195-216, 1990.

[14] M. Silva, "Petri Nets and Flexible Manufacturing," in Advances in Petri Nets 1989, G. Rozenberg (ed.), Lecture Notes of Computer Science, Springer Verlag, 1989

ISSN 0249-6399