A polynomial algorithm for a bi-criteria cyclic scheduling problem

Olivier Marchetti, Alix Munier-Kordon
Laboratoire LIP6
Université Pierre et Marie Curie
4 place Jussieu, 75 252 Paris, Cedex 05, France.
e-mail: Olivier.Marchetti@lip6.fr, Alix.Munier@lip6.fr
Fax : 33 1 44 27 72 80

Abstract

The minimization of the amount of initial tokens in a Weighted Timed Event Graph (in short WTEG) under throughput constraint is a crucial problem in industrial area such as the design of manufacturing systems or embedded systems. In this paper, an efficient polynomial algorithm is developed for the minimization of the overall places capacities with a maximum throughput. It is proved optimal for a particular sub-classes of WTEG and is leading to a 2-approximation solution in the general case.

Introduction

Cyclic scheduling problems, in which a set of generic tasks have to be performed infinitely often, have numerous practical applications. In these systems, the throughput is usually an important performance measure for designers (for surveys, see (Hanen & Munier 1994; Middendorf & Timkovsky 2002)).

In this paper, we consider that the constraints on tasks are modelled using Weighted Timed Event Graph (in short WTEG) $\mathcal{G}$ which is a subclass of Petri Nets. Transitions are associated with generic tasks and their firings have a given duration. Each place p has exactly one input and one output transition weighted by respective values $w(p)$ and $v(p)$: at the completion of a firing of the input transition of p, $w(p)$ tokens are added to p. At the firing of the output transition of p, $v(p)$ tokens are removed from p. If $v(p) = w(p) = 1$ for every place, $\mathcal{G}$ is a Timed Event Graph (in short TEG).

WTEG and TEG are widely used for modelling and solving practical cyclic scheduling problems. In the context of manufacturing systems, they are considered to model complex assembly lines. Workshop (*resp.* products) are usually modelled by transitions (*resp.* tokens). Between two successive transformations, products (*i.e.* tokens) have to be stored or have to be moved from a workshop to an other one. The amount of products, also called Work In Process (WIP in short), that have to be stored or moved may have economical consequences. Therefore, the main problem for designers is to devise an initial configuration of WIP that allows the system to reach a given productivity and that uses the smallest amount of WIP. Many optimization algorithms and heuristics were developed in order to solved it and some variants for (non weighted) Timed Event Graphs or Weighted Timed Event Graphs (*see.* (Di Febbraro, Minciardi, & Sacone 1997; Gaubert 1990; Giua, Piccaluga, & Seatzu 2002; Laftit, Proth, & Xie 1989; Sauer 2003)).

Synchronous Data-Flow (Lee & Messerschmitt 1987) is a well-known formalism considered for modelling some embedded applications and is equivalent to WTEG. In this case, transitions represent processes and places intermediate buffers. Tokens model data exchanged between processes. The size of the intermediate buffers must be minimum because of the high cost of the memories. This criteria can be expressed by associating a backwards place $p' = (t_j, t_i)$ to any place $p = (t_i, t_j)$ (*see.* (Marchetti & Munier-Kordon 2005)). The total size of the buffer corresponding to p is then the amount of tokens in places p and p'. The bi-criteria problem studied is then to minimize the total number of initial tokens under throughput constraint. Researchers have addressed some closely related problems by various approaches such as integer linear programming (Govindarajan, Gao, & Desai 2002) or more recently by model checking (Stuijk, Geilen, & Basten 2006).

This paper is dedicated to the presentation of a polynomial algorithm to solve the minimization of the overall places capacities to achieve a maximum throughput. The following section is dedicated to basic definitions and the description of our problem. In Section 3, the algorithm and its performance are presented. It is proved that it leads to an optimal solution for particular durations of the firings. In the general case, a 2-approximation solution is obtained. We conclude lastly with some perspectives.

Notations and definitions

Basic definitions and examples

We suppose that the reader is aware with theoretical background of Petri Nets (*see.* (Peterson 1981) for further details). However, this subsection recalls the main definitions concerning this paper.

Weighted Event Graph A Weighted Event Graph $\mathcal{G} = (P, T)$ (in short WEG) is given by a set of places $P = \{p_1, \ldots, p_m\}$ and a set of transitions $T = \{t_1, \ldots, t_n\}$. Every place $p \in P$ is defined between two transitions t_i and t_j

and is denoted by $p = (t_i, t_j)$. The arcs (t_i, p) and (p, t_j) are valued by strictly positive integers denoted respectively by $w(p)$ and $v(p)$ (see Figure 1). At each firing of the transition t_i (*resp.* t_j), $w(p)$ (*resp.* $v(p)$) tokens are added to (*resp.* removed from) place p. If $v(p) = w(p) = 1$ for every place $p \in P$, then $\mathcal{G}$ is an Event Graph (in short EG).

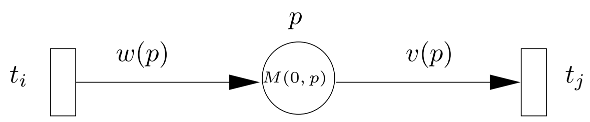

Figure 1: A place $p = (t_i, t_j)$ of a Generalized Timed Event Graph.

We will assume that there is at most one place $p = (t_i, t_j)$ defined from t_i to t_j. For any transition $t \in T$, we set

$$\mathcal{P}^+(t) = \{p = (t, t') \in P, t' \in T\}$$

and

$$\mathcal{P}^-(t) = \{p = (t', t) \in P, t' \in T\}$$

For any place $p \in P$, gcd_p and lcm_p denotes respectively the greatest common divisor and the least common multiple of integers $v(p)$ and $w(p)$.

The incidence matrix I associated with a WEG $\mathcal{G}$ is defined by $|P| \times |T|$ values in $\mathbb{Z}$ such that, for any couple $(p, t) \in P \times T$:

- $I[p, t] = w(p)$ (*resp.* $-v(p)$) if $p \in \mathcal{P}^+(t)$ (*resp.* if $p \in \mathcal{P}^-(t)$) ;
- $I[p, t] = 0$ otherwise.

A path μ of $\mathcal{G}$ is defined as a sequence of α places such that $\mu = (p_1 = (t_1, t_2), p_2 = (t_2, t_3), \ldots, p_\alpha = (t_\alpha, t_{\alpha+1}))$. If this path is closed (*i.e.* $t_1 = t_{\alpha+1}$), then μ is a circuit. An initial marking is a function $M_0 : P \rightarrow \mathbb{N}$ such that, for any $p \in P$, $M_0(p)$ denotes the initial number of tokens in the place p. A WEG is marked if it is associated with an initial marking.

Timed Weighted Event Graph A Timed Weighted Event Graph (in short TWEG) is a WEG associated with a function $\ell : T \rightarrow \mathbb{R}^{+\star}$ such that, for any $t \in T$, $\ell(t)$ is the duration of a firing of t. It is usually denoted by $\mathcal{G} = (P, T, \ell)$. For every place $p = (t_i, t_j) \in P$, $w(p)$ tokens are added to p exactly $\ell(t_i)$ times units after the firing of t_i. $v(p)$ tokens are removed from p at the firing of t_j.

We suppose that two successive firings of the same transition cannot overlap: this is modelled by a self loop place $p = (t_i, t_i)$, $\forall t_i \in T$ with $w(p) = v(p) = 1$ and $M_0(p) = 1$. For a sake of simplicity, these loops are not presented in figures.

We denote by $M(\tau, p)$ the instantaneous marking of the place p at time instant $\tau \geq 0$. The marking $M(0, p)$ is called the initial marking of place p and $M(\mathcal{G})$ points out the initial marking of the WTEG $\mathcal{G}$. $\mathcal{G}$ is a marked WTEG (*resp.* TEG) if it is associated with an initial marking $M(\mathcal{G})$.

Example 1 A Digital Signal Processing (DSP in short) may be modelled by a TWEG: transitions represent processes that may be executed infinitely often. Buffers allow communications between processes. A buffer has exactly one input and one output and data stored in a buffer have the same size. At the beginning (*resp.* ending) of each iteration, a process has to read (*resp.* write) a fixed amount of data in each input (*resp.* output) buffers. Let us suppose that the WTEG pictured by Figure 2 depicts a DSP. Transition t_3 needs 2 data from t_1 and 7 data from t_2 to be executed once. At its completion, 5 data are temporarily stored in place $p_3 = (t_3, t_4)$ to be sent to t_4.

The completion of one execution of each process t_i has a duration $\ell(t_i)$. Because of the cost of the memories in embedded systems, the capacity of the buffers (*i.e.* the maximum number of data that can be simultaneously stored) should be limited. So, the problem consists of finding an initial marking of the places and their capacities $C(p_i)$, $i \in \{1, \ldots, m\}$ such that several criteria are realized:

- The system should be live: each transition $t \in T$ may be fired infinitely often.
- The surface of a buffer on the chip is usually proportional to the size of the data stored in it. More formally, if y_p, $p \in P$ denotes the size of one data stored in p, the criteria considered is the minimization of the global storage surface defined as $\sum_{p \in P} y_p \cdot C(p)$.
- The throughput of the system should be maximum.

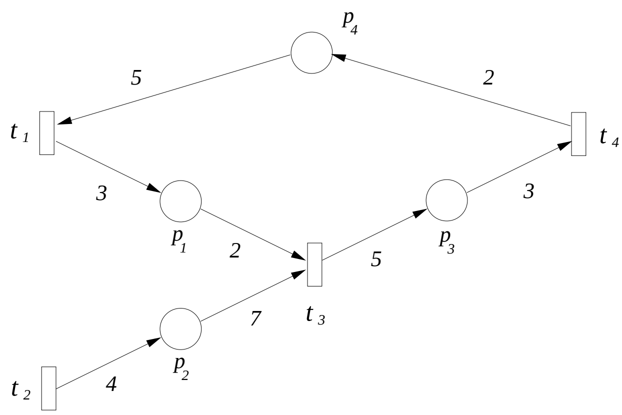

Figure 2: A Weighted Timed Event Graph with $\ell(t_1) = 5$, $\ell(t_2) = 2$, $\ell(t_3) = 6$ and $\ell(t_4) = 3$.

Example 2 Let us consider the assembling of products P_1 and P_2 from raw materials M_1, M_2 and M_3 following three levels as pictured by Figure 3. Level 0 corresponds to the final assembling, level 2 to the loading of material raws on the line. It is also assumed that a product at level $l > 0$ may be used for only one operation at level $l - 1$. Each arc (i, j) is valued by an integer corresponding to the number of products i needed to get one product j.

Moreover, each workshop t_i is dedicated to exactly one operation (*i.e.* there is no conflict between assembling operations) and is composed by one machine (*i.e* two distinct

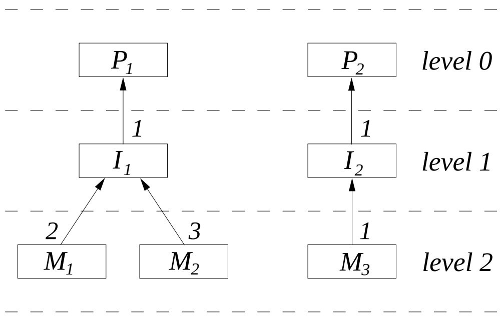

Figure 3: Assembling operations.

products cannot be assembled simultaneously by the same workshop). Operations and their corresponding durations are given by the following table:

workshop	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8
operations	M_1	M_2	M_3	I_1	I_2	P_1	P_2	$\cdot$
durations	2	2	2	15	12	12	6	14

A transporter takes 3 products P_1 and 2 products P_2 and brings 8 raw materials to M_1, 12 to M_2 and 2 to M_3.

A modelling of this assembly line using a TWEG is presented by Figure 4.

However, each product has to be stored between two successive operations. Storing products between workshops has economical consequences and therefore it is important to limit the whole storage surface in the assembly line. Again, the problem consists of finding an initial marking of the places and their capacities $C(p_i)$, $i \in \{1, \ldots, m\}$ such that several criteria are realized:

- The assembly line should be live: each transition $t \in T$ may be fired infinitely often.
- Economical costs due to the storage problem are modeled by an integer value y_p for each place p. The criteria considered is the minimization of the global storage weighted surface defined as $\sum_{p \in P} y_p \cdot C(p)$.
- The throughput of the assembly line should be maximum.

In the following, we recall some results concerning these three criteria.

Weight of paths and Liveness of a WEG

The complexity of the determination of the liveness of a marked WEG is unknown. Algorithms developed up to now to answer this question are pseudo-polynomial (*see.* as example (Munier 1993)).

However, it is polynomial for marked Event Graphs. Indeed, setting $H(\mathcal{C}) = \sum_{p \in P \cap \mathcal{C}} M(0,p)$ the height of a circuit $\mathcal{C}$ of $\mathcal{G}$, it is proved in (Commoner *et al.* 1971) that M is a live marking iff the height of every circuit of $\mathcal{G}$ is strictly positive.

In the case of Weighted Event Graph, the following simple necessary condition of liveness was noticed by several authors (Karp & Miller 1966; Munier 1993; Teruel *et al.* 1992): let us define the weight (or gain) of every path μ of a

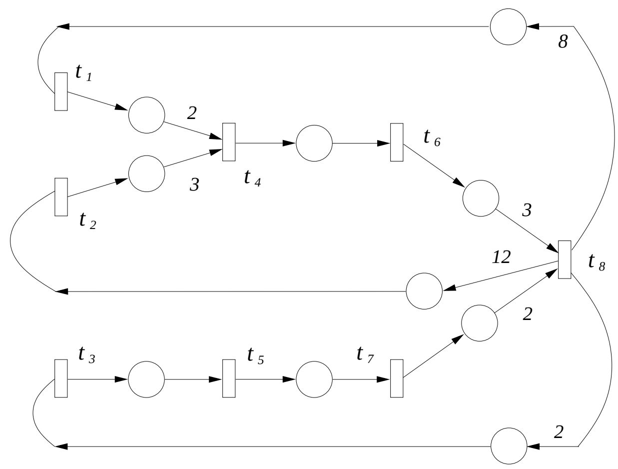

Figure 4: Modelling of the assembly line using a TWEG. $\forall t_i$, $\ell(t_i)$ is the duration of the corresponding operation.

Weighted Event Graph $\mathcal{G}$, denoted by $W(\mu)$ as

$$W(\mu) = \prod_{p \in P \cap \mu} \frac{w(p)}{v(p)}.$$

Then, if a marked Weighted Event Graph $\mathcal{G}$ is live, every circuit has a weight equal to or greater than 1. This condition is trivially not sufficient: it is always fulfilled for Event Graphs and not sufficient to decide about the liveness.

Symmetric Weighted Event Graphs

One of the criteria considered here is the minimization of a linear function of the total capacity of the places. A bounded capacity $C(p)$ of a place $p = (t_i, t_j) \in P$ may be modelled by adding another place $p' = (t_j, t_i)$ with $w(p') = v(p)$, $v(p') = w(p)$ and such that the initial marking verifies $M_0(p) + M_0(p') = C(p)$. So, a Symmetric WEG defined as follows may be associated to any WEG with bounded capacity places:

Proposition 1. *Let $\mathcal{G}$ be a WEG:*

1. *$\mathcal{G}$ is a Symmetric Weighted Event Graph (SWEG in short) if each place $p = (t_i, t_j) \in P$ is associated with a backward place $p' = (t_j, t_i) \in P$ with $w(p') = v(p)$ and $v(p') = w(p)$. (p, p') is then called a couple of backward places.*
2. *An initial marking $M(\mathcal{G})$ of a SWEG $\mathcal{G}$ is said minimally bounded if for any couple (p, p') of backward places,*

$$M_0(p) + M_0(p') = M_{min}(p) = w(p) + v(p) - gcd_p.$$

Note that $M_{min}(p) = 1$, $\forall p \in P$ if $\mathcal{G}$ is a Symmetric Event Graph.

The second part of the definition comes from (Marchetti & Munier-Kordon 2004): it is proved that the whole number of tokens of any couple of backward places (p, p') of a WEG is constant and must be greater than or equal to $M_{min}(p)$ if the graph is live.

For our practical problem, we may considered a SWEG $\mathcal{G}' = (P', T, \ell)$ built from a WEG $\mathcal{G} = (P, T, \ell)$ by adding a backward place to any place $p \in P$. The liveness problem consists then to find an initial live marking $M(0,p), p \in P'$. The places capacities are bounded and verify, for any $p \in P$ and its backward place p', $C(p) = M(0,p) + M(0,p')$. So, $M_{min}(p), p \in P$ is a lower bound of $C(p)$.

Unitary graphs and Normalization

Unitary graphs Unitary WEG were introduced in (Munier 1993) as follows:

Definition 1. *A unitary WEG $\mathcal{G}$ is strongly connected and such that every circuit of $\mathcal{G}$ has a unit weight.*

The SWEG tackled in this paper are unitary graphs: they are strongly connected. Moreover, if there exists a circuit c with $W(c) > 1$, then a circuit c' may be built by considering backward places with $W(c') < 1$. By the necessary condition of liveness expressed previously, we conclude that the SWEG is not live.

Normalization of a unitary graph Normalization of a unitary marked WEG $\mathcal{G}$ is a transformation of all the marking functions and initial marking values such that the marking functions adjacent to any transition t_i have the same value Z_i. This transformation does not affect the constraints induced by places between the firings, so that these two graphs are equivalent. More formally:

Definition 2. *A transition t_i is normalized iff there exists $Z_i \in \mathbb{N}^\star$ such that:*

$$\begin{array}{lll} \forall p \in \mathcal{P}^+(t_i), & w(p) & = Z_i \text{ and} \\ \forall p \in \mathcal{P}^-(t_i), & v(p) & = Z_i. \end{array}$$

A WEG G is said to be normalized iff all its transitions are normalized.

In (Marchetti & Munier-Kordon 2004), it is stated that any unitary Weighted Event Graph can be polynomially transformed into an equivalent normalized Weighted Event Graph by modifying marking functions and initial markings. So, we consider in this paper normalized SWEG.

Example Let us consider the WTEG pictured by Figure 2. The corresponding SWTEG is pictured by Figure 5. Notice that it is a unitary graph.

The normalization step consists in solving the following system:

$$\begin{array}{lll} \forall p \in P, & x(p) \in \mathbb{N}^\star & \\ \forall t_i \in T, & Z_i \in \mathbb{N}^\star & \\ \forall p \in \mathcal{P}^+(t_i), & w(p) \cdot x(p) & = Z_i \\ \forall p \in \mathcal{P}^-(t_i), & v(p) \cdot x(p) & = Z_i. \end{array}$$

It is proved in (Marchetti & Munier-Kordon 2004) that a solution exists. For our example, equations are $Z_1 = 3 \cdot x(p_1) = 5 \cdot x(p_4)$, $Z_3 = 7 \cdot x(p_2) = 2 \cdot x(p_1) = 5 \cdot x(p_3)$ and $Z_4 = 3 \cdot x(p_3) = 2 \cdot x(p_4)$. A solution is $x(p_1) = 35$, $x(p_2) = 10$, $x(p_3) = 14$, $x(p_4) = 21$ and $(Z_1, Z_2, Z_3, Z_4) = (105, 40, 70, 42)$. The corresponding normalized SWEG is pictured by Figure 6.

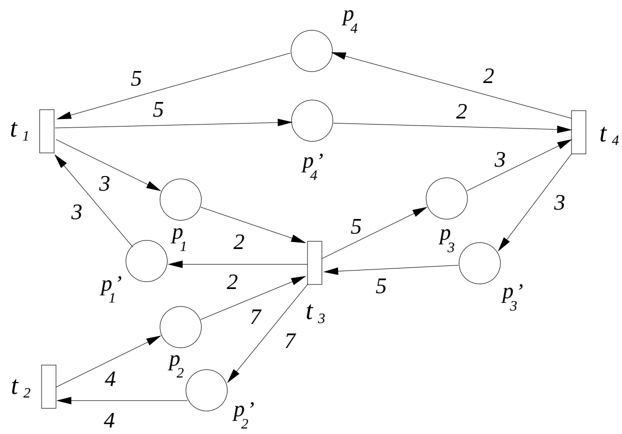

Figure 5: The Symmetric Timed Event Graph associated with the WTEG of Figure 2.

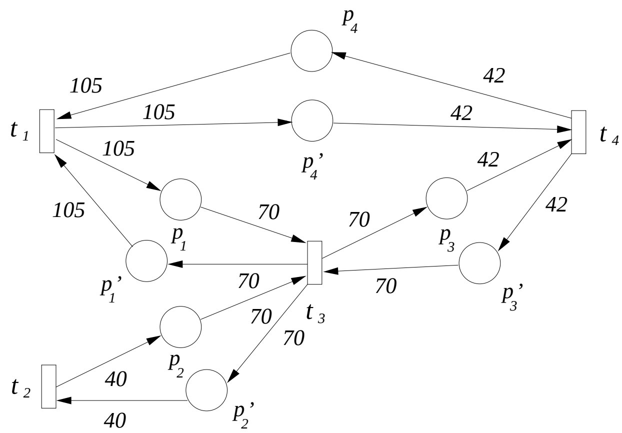

Figure 6: The Symmetric Normalized Timed Event Graph associated with the SWTEG of Figure 5.

Schedules and throughput of unitary normalized WTEGs

Let $\mathcal{G}$ be a unitary marked WTEG. A schedule is a function $\sigma : T \times \mathbb{N}^\star \to \mathbb{R}^{+\star}$ such that $\sigma(t_i, q)$ denotes the starting time of the q-th firing of t_i. There is a strong relationship between a schedule and the corresponding instantaneous marking. Let $p = (t_i, t_j)$ be a place of P. For any value $\tau \in \mathbb{R}^{+\star}$, let us denote by $E(\tau, t_i)$ the number of firings of t_i completed at time τ. More formally,

$$E(\tau, t_i) = \max\{q \in \mathbb{N}, \sigma(t_i, q) + \ell(t_i) \leq \tau\}.$$

On the same way, $B(\tau, t_j)$ denotes the number of firings of t_j at time τ and

$$B(\tau, t_j) = \max\{q \in \mathbb{N}, \sigma(t_j, q) \leq \tau\}.$$

Clearly,

$$M(\tau, p) = M(0, p) + w(p) \cdot E(\tau, t_i) - v(p) \cdot B(\tau, t_j).$$

A schedule (and its corresponding marking) is feasible if $M(\tau, p) \geq 0$ for every couple $(\tau, p) \in \mathbb{R}^{+\star} \times P$. The throughput of a transition t_i for a schedule σ is defined as

$$\lambda(t_i) = \lim_{q\to\infty} \frac{q}{\sigma(t_i, q)}.$$

Note that for live marked WTEG, earliest schedule (which consists on firing the transitions as soon as possible) always exists and has maximum throughputs. We consider here that the throughputs of a marked WTEG are those of its earliest schedule.

Definition 3. *A sequence u_n, $n \in \mathbb{N}^\star$ is K-periodic if there exists $(N_0, K, w) \in \mathbb{N}^2 \times \mathbb{R}^{+\star}$ such that, for any $n \geq N_0$, $u_{n+K} = u_n + w$. If $K = 1$, the sequence is said to be periodic.*
*$\frac{K}{w}$ is the throughput of the sequence and K its periodicity factor. A schedule σ is K-periodic (*resp.* periodic) if sequences $\sigma(t_i, k), k \in \mathbb{N}$ are K-periodic (*resp.* periodic), for any $t_i \in T$.*

For marked TEGs, the determination of the throughput of the earliest can be done polynomially (Chrétienne 1983). The computation of the throughputs has an unknown complexity for marked WTEGs. Up to now, the algorithms developed are pseudo-polynomial time (*see.* as example (Munier 1993)). However, we deduce from (Munier 1993) and the definition of Z_i, $i \in \{1, \ldots, n\}$ the following theorem:

Theorem 1. *Let $\mathcal{G}$ be a unitary WTEG with a live initial marking $M(\mathcal{G})$. Then, the earliest schedule is K-periodic and its throughputs verify, for any couple of transitions $(t_i, t_j) \in T^2$,*

$$\lambda(t_i) \cdot Z_i = \lambda(t_j) \cdot Z_j \stackrel{def}{=} \lambda(M(\mathcal{G})).$$

$\lambda(M(\mathcal{G}))$ is called the throughput of $\mathcal{G}$ associated with the initial marking $M(\mathcal{G})$.

Let us suppose now that the initial marking $M(\mathcal{G})$ of a strongly connected WTEG is such that, for the earliest schedule, at least one transition $t_{i^\star}$ is fired without any interruption. Then, $\lambda(t_{i^\star}) = \frac{1}{\ell(t_{i^\star})}$. By Theorem 1 and since $\lambda(t_i) \leq \frac{1}{\ell(t_i)}$ for any transition $t_i \in T$, we obtain:

$$\lambda(M(\mathcal{G})) = \frac{Z_{i^\star}}{\ell(t_{i^\star})} \leq \frac{Z_i}{\ell(t_i)}, \forall t_i \in T.$$

We deduce the following definition:

Definition 4. *The intrinsic throughput of a unitary WTEG $\mathcal{G}$ is defined as the ratio $\min_{t_i\in T}\left(\frac{Z_i}{\ell(t_i)}\right)$ and corresponds to the maximum throughput that may be achieved by an initial marking of $\mathcal{G}$.*

This value may always be achieved if we consider a sufficiently large number of tokens in each place.

Let us consider our previous example depicted by Figure 6. Then $\frac{Z_1}{\ell(t_1)} = \frac{105}{5} = 21$, $\frac{Z_2}{\ell(t_2)} = \frac{40}{2} = 20$, $\frac{Z_3}{\ell(t_3)} = \frac{70}{6} = \frac{35}{3}$ and $\frac{Z_4}{\ell(t_4)} = \frac{42}{3}$. So $t_{i^\star} = t_3$ and the intrinsic throughput of the SWTEG pictured by Figure 6 is $\frac{35}{3}$.

Problem formulation

The problem tackled here may be formulated as follows: MAX INTRINSIC THROUGHPUT:

Instance: $\mathcal{G} = (P, T, \ell)$ is a normalized SWTEG, a vector y_p, $p \in P$ such that $y_p = y_{p'}$ for every couple (p, p') of backward places.

Question: Is there an initial live marking $M(\mathcal{G})$ such that $\sum_{p\in P} y_p \cdot M(0, p)$ is minimized for a throughput and $\lambda(M(\mathcal{G})) = \min_{t_i\in T}\left(\frac{Z_i}{\ell(t_i)}\right)$?

The complexity of the corresponding decision problem was proved NP-complete in (Marchetti & Munier-Kordon 2006).

A polynomial time algorithm

We first consider a couple of backward places (p_1, p_2) between two transitions t_i and t_j with $p_1 = (t_i, t_j)$, $v(p_1) = w(p_2) = Z_j$, $v(p_2) = w(p_1) = Z_i$ and $\frac{Z_i}{\ell(t_i)} = \frac{Z_j}{\ell(t_j)} = \rho$. This SWTEG is denoted by $\mathcal{G}_{p_1,p_2}$ (*see.* Figure 7).

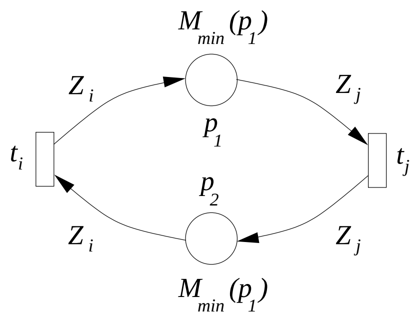

Figure 7: A SWTEG $\mathcal{G}_{p_1,p_2}$ marked with $M(0, p_1) = M(0, p_2) = M_{min}(p_1)$.

The aim of this section is to prove that, setting $M(0, p_1) = M(0, p_2) = M_{min}(p_1)$ is an optimal solution for the MAX INTRINSIC THROUGHPUT problem for this graph. We also show that this initial marking leads to a 2-approximation solution for general SWTEG.

We first show that this marking leads to an optimal periodic schedule for G_{p_1,p_2}. Then, by studying the earliest schedule of $\mathcal{G}_{p_1,p_2}$ for a schedule with a maximum intrinsic throughput, we prove that $2M_{min}(p_1)$ is an upper bound on the number of tokens for $\mathcal{G}_{p_1,p_2}$. We conclude that this marking is optimal for $\mathcal{G}_{p_1,p_2}$ and for any unitary SWEGT such that $\ell(t_i) = \rho.Z_i$ for any transition $t_i \in T$. We prove then that, for any unitary SWEGT with no assumption on the durations of the firing, the previous simple algorithm leads to a 2-approximation solution.

Feasibility of the initial marking for $\mathcal{G}_{p_1,p_2}$

Lemma 1. *The schedule $\sigma(t, k) = (k - 1) \cdot \ell(t)$ for $t \in \{t_i, t_j\}$ is feasible for the initial marking $M(0, p_1) = M(0, p_2) = M_{min}(p_1)$ and its throughput equals the intrinsic throughput of $\mathcal{G}_{p_1,p_2}$.*

Proof. By definition of σ, for any $p = (t, t')$, $t \neq t'$ with $(t, t') \in \{t_i, t_j\}^2$ and $\tau \in \mathbb{R}$, the number of firings of t completed at time τ is $E(\tau, t) = \left\lfloor \frac{\tau}{\ell(t)} \right\rfloor = \left\lfloor \frac{\tau}{\rho \cdot w(p)} \right\rfloor$. In the same way, the number of firings of t' at time τ is $B(\tau, t') = \left\lceil \frac{\tau}{\ell(t')} \right\rceil = \left\lceil \frac{\tau}{\rho \cdot v(p)} \right\rceil$. So, the instantaneous marking of p can be described by the following equations:

$$M(0, p) = w(p) + v(p) - gcd_p$$

For any $\tau > 0$,

$$M(\tau, p) = M(0, p) + \left\lfloor \frac{\tau}{\rho \cdot w(p)} \right\rfloor \cdot w(p) - \left\lceil \frac{\tau}{\rho \cdot v(p)} \right\rceil \cdot v(p)$$

We prove that $M(\tau, p) \geq 0$ for any value $\tau \geq 0$ by contradiction. Assume that there exists a value $\tau \geq 0$ such that $M(\tau, p) < 0$. Since $M(\tau, p)$ is a linear combination of $w(p)$ and $v(p)$ it follows that $M(\tau, p) \equiv 0 \mod gcd_p$ and therefore

$$M(\tau, p) < 0 \Longleftrightarrow M(\tau, p) \leqslant -gcd_p.$$

So, by definition of $M(\tau, p)$, we obtain

$$M(0, p) + \left\lfloor \frac{\tau}{\rho \cdot w(p)} \right\rfloor \cdot w(p) - \left\lceil \frac{\tau}{\rho \cdot v(p)} \right\rceil \cdot v(p) \leqslant -gcd_p.$$

Since $M(0, p) = M_{min}(p) = w(p) + v(p) - gcd_p$, we deduce that

$$w(p) + v(p) + \left\lfloor \frac{\tau}{\rho \cdot w(p)} \right\rfloor \cdot w(p) \leqslant \left\lceil \frac{\tau}{\rho \cdot v(p)} \right\rceil \cdot v(p)$$

However $\forall x \in \mathbb{R}$ we have

$$x - 1 < \lfloor x \rfloor \leqslant x \leqslant \lceil x \rceil < x + 1.$$

Therefore

$$w(p) + v(p) + \frac{\tau}{\rho} - w(p) < \frac{\tau}{\rho} + v(p)$$

and then

$$0 < 0.$$

A contradiction. So, the schedule is feasible. Moreover, we have $\lambda(t) = \frac{1}{\ell(t)}$ for $t \in \{t_i, t_j\}$. As $\rho = \frac{Z_i}{\ell(t_i)} = \frac{Z_j}{\ell(t_j)}$, we deduce by Theorem 1 that the throughput of σ equals ρ and is exactly the intrinsic throughput of $\mathcal{G}_{p_1,p_2}$. □

Earliest schedule of $\mathcal{G}_{p_1,p_2}$

We suppose here that $\mathcal{G}_{p_1,p_2}$ has an initial marking $M(\mathcal{G})$ which values are not necessarily equal to $M_{min}(p_1)$.

Lemma 2. *If the throughput of the earliest schedule s of $\mathcal{G}_{p_1,p_2}$ equals ρ for a given marking, then s is periodic. Moreover, it exists $\tau^\star \in \mathbb{R}^+$ such that, for any $t \in \{t_i, t_j\}$ there exists $k_t \in \mathbb{N}$ such that $s(t, k) = \tau^\star + (k - k_t) \cdot \ell(t)$.*

Proof. By Theorem 1, the earliest schedule s is K-periodic. Now, since the system throughput is intrinsically maximum, there is no idle times when time τ tends to infinity and the periodicity factor K of the sequences is unit.

So, for any $t \in \{t_i, t_j\}$, there exists a couple $(\tau^\star(t), k(t)) \in \mathbb{R} \times \mathbb{N}^\star$ with minimum values such that, for any $k \geq k(t)$, $s(t, k) = \tau^\star(t) + (k - k(t))\ell(t)$.

If $\tau^\star(t_j) = \tau^\star(t_i)$, the lemma is proved by setting $\tau^\star = \tau^\star(t_j)$ and $k_t = k(t)$ for $t \in \{t_i, t_j\}$.

Let us suppose now that $\tau^\star(t_j) > \tau^\star(t_i)$. Then, by minimality of $\tau^\star(t_j)$, there is an idle slot before $s(t_j, k(t_j))$ due to a precedence constraint before the $k(t_j)$ firing of t_j. So, $s(t_j, k(t_j))$ is the completion time of a firing of t_i and there exists an integer k' such that $s(t_i, k') = s(t_j, k(t_j))$. Setting $\tau^\star = \tau^\star(t_j)$, $k_{t_i} = k'$ and $k_{t_j} = k(t_j)$, we obtain the result.

Lastly, if $\tau^\star(t_i) > \tau^\star(t_j)$, the same reasoning holds by swapping i and j.

□

In order to show that our solution is minimum, we have to study the marking of places of $\mathcal{G}_{p_1,p_2}$ whenever both transitions t_i and t_j are executed at the same instant time. In this purpose, we define a special period of time.

Let us now define N_i and N_j as the smallest strictly positive integers such that $N_i \cdot w(p_1) = N_j \cdot v(p_1)$. Observe that, since $w(p_1) = Z_i$ and $v(p_1) = Z_j$, $N_i \cdot Z_i = N_j \cdot Z_j$ and $\dfrac{\ell(t_i)}{\ell(t_j)} = \dfrac{Z_i}{Z_j} = \dfrac{N_j}{N_i}$. We denote $H = \ell(t_i) \cdot N_i = \ell(t_j) \cdot N_j$ the hyperperiod of the graph $\mathcal{G}_{p_1,p_2}$. This value is associated with the earliest schedule of $\mathcal{G}_{p_1,p_2}$. Therefore, by lemma 2, for each date $\tau > \tau^\star$ the instantaneous marking $M(\tau, p)$ of place $p \in \{p_1, p_2\}$ is defined by:

$$M(\tau, p) = M(\tau^\star, p) + \left\lfloor \frac{\tau - \tau^\star}{\rho \cdot w(p)} \right\rfloor \cdot w(p) - \left\lceil \frac{\tau - \tau^\star}{\rho \cdot v(p)} \right\rceil \cdot v(p)$$

Then, for each date $\tau > \tau^\star$, we have:

$$M(\tau + H, p) = M(\tau^\star, p) + \left\lfloor \frac{\tau + H - \tau^\star}{l(t_i)} \right\rfloor \cdot w(p) - \left\lceil \frac{\tau + H - \tau^\star}{l(t_j)} \right\rceil \cdot v(p)$$

$$M(\tau + H, p) = M(\tau^\star, p) + \left(\left\lfloor \frac{\tau - \tau^\star}{l(t_i)} \right\rfloor + N_i \right) \cdot w(p) - \left(\left\lceil \frac{\tau - \tau^\star}{l(t_j)} \right\rceil + N_j \right) \cdot v(p)$$

$$M(\tau + H, p) = M(\tau^\star, p) + \left\lfloor \frac{\tau - \tau^\star}{l(t_i)} \right\rfloor \cdot w(p) - \left\lceil \frac{\tau - \tau^\star}{l(t_j)} \right\rceil \cdot v(p)$$

$$M(\tau + H, p) = M(\tau, p)$$

So, we can reduce the study of marking to a period of time that lasts an hyperperiod H.

Now, observe that the time interval $[\tau^\star, \tau^\star + H[$ contains exactly N_i firings of t_i and N_j firings of t_j such that:

- t_i and t_j are fired simultaneously at time $\tau^\star$,
- by minimality of N_i and N_j, there are exactly $N_i - 1$ firings of t_i and $N_j - 1$ firing of t_j in the interval $]\tau^\star, \tau^\star + H[$ and they cannot be simultaneously (otherwise N_i and N_j would not be minimum).

So, a strictly increasing sequence of real number $s_h, h \in \{0, \ldots, N_i + N_j - 1\}$ can be defined such that $s_0 = \tau^\star$, $s_{N_i+N_j-1} = \tau^\star + H$ and, for any $r \in \{1, \ldots, N_i + N_j - 2\}$, a firing of t_i or t_j occurs at time s_r. We also define the interval sequence C_r, $r \in \{1, \ldots, N_i + N_j - 1\}$ as $C_r =]s_{r-1}, s_r[$.

For example, let us consider the couple of backwards places pictured by Figure 8. We assume that the initial marking (not presented in Figure) can not prevent any firing of transition t_i and t_j. We have $N_i = 5$ and $N_i = 2$. Then $H = 20$ and a slice of the earliest schedule between $\tau^\star$ and $\tau^\star + H$ is pictured by Figure 9. We get $s_0 = \tau^\star$, $s_1 = \tau^\star + 4$, $s_2 = \tau^\star + 8$, $s_3 = \tau^\star + 10$, $s_4 = \tau^\star + 12$, $s_5 = \tau^\star + 16$ and $s_6 = \tau^\star + 20$.

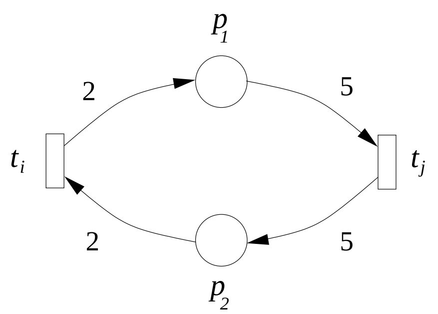

Figure 8: A couple of backward place with $\ell(t_i) = 4$ and $\ell(t_j) = 10$. Notice that $\frac{\ell(t_i)}{Z_i} = \frac{\ell(t_j)}{Z_j} = 2$.

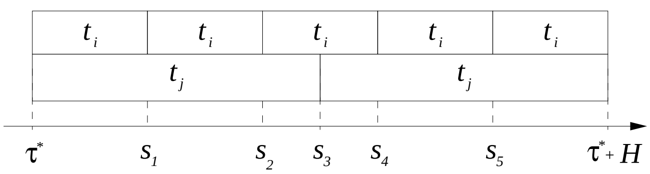

Figure 9: The earliest schedule between $\tau^\star$ and $\tau^\star + H$.

Now, for any $r \in \{1, \ldots, N_i + N_j - 2\}$, u_r denotes the transition of $\mathcal{G}_{p_1,p_2}$ fired at time s_r. $u = u_1 \ldots u_{N_i+N_j-2}$ is then the firing sequence of $\mathcal{G}_{p_1,p_2}$ during the time interval $]\tau^\star, \tau^\star + H[$.

For our last example, we get $u_1 = t_i$, $u_2 = t_i$, $u_3 = t_j$, $u_4 = t_i$ and $u_5 = t_i$.

We build now a marked Weighted Event Graph (not timed) denoted by $\mathcal{G}'_{p_1,p_2}$ as follows:

1. $\mathcal{G}'_{p_1,p_2}$ has the same structure as $\mathcal{G}_{p_1,p_2}$ (same places and same transitions).
2. Its initial marking is $M'_0(p) = M(\tau^\star, p) - v(p)$ for $p \in \{p_1, p_2\}$.

Lemma 3. *$u = u_1 \ldots u_{N_i+N_j-2}$ is a firing sequence of $\mathcal{G}'_{p_1,p_2}$.*

Proof. Notice that, for any couple $(\tau_1, \tau_2) \in C_r^2$, $r \in \{1, \ldots, N_i + N_j - 1\}$, $M(\tau_1, p) = M(\tau_2, p)$ for $p \in \{p_1, p_2\}$. So, the marking function M of $\mathcal{G}_{p_1,p_2}$ is constant on C_r, $r \in \{1, \ldots, N_i + N_j - 1\}$.

For any $r \in \{1, \ldots, N_i + N_j - 1\}$, we prove by recurrence that the marking of places of $\mathcal{G}'_{p_1,p_2}$ after the sequence of firings $u_1 \ldots u_{r-1}$ is exactly $M(\tau, p)$, $p \in \{p_1, p_2\}$ and $\tau \in C_r$.

1. For any $\tau \in C_1$, $M(\tau, p_1) = M(\tau^\star, p_1) - v(p_1)$ and $M(\tau, p_2) = M(\tau^\star, p_2) - v(p_2)$. So it is true for $r = 1$.
2. Now, let us suppose that it is true for a value $1 \le r < N_i + N_j - 1$. Let us consider a couple of times $(\tau_1, \tau_2) \in C_r \times C_{r+1}$.
   - If $u_r = t_i$, then we get $M(\tau_2, p_1) = M(\tau_1, p_1) + w(p_1)$ and $M(\tau_2, p_2) = M(\tau_1, p_2) - w(p_1)$. By hypothesis, the markings of p_1 and p_2 after the sequence $u_1 \ldots u_{r-1}$ are respectively equal to $M'(p_1) = M(\tau_1, p_1)$ and $M'(p_2) = M(\tau_1, p_2)$. So, after the firing of $u_r = t_i$, they are respectively equal to $M(\tau_2, p_1)$ and $M(\tau_2, p_2)$.
   - In the same way, if $u_r = t_j$, we obtain $M(\tau_2, p_1) = M(\tau_1, p_1) - w(p_2)$ and $M(\tau_2, p_2) = M(\tau_1, p_2) + w(p_2)$. The property is also true in this case.

So, we conclude that $u = u_1 \ldots u_{N_i+N_j-2}$ is a firing sequence of $\mathcal{G}'_{p_1,p_2}$. □

A state of $\mathcal{G}'_{p_1,p_2}$ is given by a couple of marking $\begin{pmatrix} M'(p_1) \\ M'(p_2) \end{pmatrix}$. Next subsection provides some properties on the number of different states of $\mathcal{G}'_{p_1,p_2}$.

Upper bound on the number of states of $\mathcal{G}'_{p_1,p_2}$

Lemma 4. *If $M'_0(p_1) + M'_0(p_2) < v(p_1) + w(p_1) - 2 \cdot gcd_{p_1}$ then all the states of $\mathcal{G}'_{p_1,p_2}$ are distinct.*

Proof. According to (Marchetti & Munier-Kordon 2004), the marked WEG $\mathcal{G}'_{p_1,p_2}$ is live iff $M'_0(p_1) + M'_0(p_2) \ge v(p_1) + w(p_1) - gcd_{p_1}$. So, it is not live and all the intermediary states are different (otherwise, the system would be live). □

Lemma 5. *If $M'_0(p_1) + M'_0(p_2) < v(p_1) + w(p_1) - 2 \cdot gcd_{p_1}$ then the number of states of $\mathcal{G}'_{p_1,p_2}$ is at most $N_i + N_j - 2$.*

Proof. By (Marchetti & Munier-Kordon 2004), we know that the number of tokens hold in a place p is a multiple of gcd_p and that the number of tokens present in both places is constant for all valid firing sequences. It follows that we can consider that $M'_0(p_1) + M'_0(p_2) \le w(p_1) + v(p_1) - 3 \cdot gcd_{p_1}$. Setting $S = v(p_1) + w(p_1) - 3 \cdot gcd_{p_1}$, the different possible states $\begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{N}^2$ can be enumerated under the following conditions:

$$\left\{\begin{array}{rcl} a & \equiv & 0 \mod (gcd_{p_1}) \\ b & \equiv & 0 \mod (gcd_{p_1}) \\ a+b & = & S \end{array}\right.$$

which yields the following enumeration

$$\underbrace{\begin{array}{c|c|c|c|c|c} 0 & gcd_{p_1} & 2\cdot gcd_{p_1} & \ldots & S-gcd_{p_1} & S \\ S & S-gcd_{p_1} & S-2\cdot gcd_{p_1} & \ldots & gcd_{p_1} & 0 \end{array}}_{\text{Couples } \left(\begin{array}{c} a \\ b \end{array}\right) \text{ with } a+b=S}$$

$$\text{This yields to } \frac{S}{gcd_{p_1}} + 1 \text{ couples}$$

By definition of N_i and N_j, we also have $N_i \cdot w(p_1) = N_j \cdot v(p_1) = lcm_{p_1}$. Since $w(p_1) \cdot v(p_1) = lcm_{p_1} \cdot gcd_{p_1}$,

$$\frac{w(p_1)}{gcd_{p_1}} = N_j \quad \text{and} \quad \frac{v(p_1)}{gcd_{p_1}} = N_i.$$

So the number of different states is equal to $N_i + N_j - 2$ for S initial tokens and lemma holds. □

Polynomial special cases and an approximation algorithm

We are now ready to show our main result:

Theorem 2. *Let $\mathcal{G}_{p_1,p_2}$ be a two backward places system such that*

$$\forall t_i \in T,\ \ \ell(t_i) = \rho \cdot Z_i\ ,\ \ \rho > 0.$$

Then setting $M(0, p_1) = M(0, p_2) = M_{min}(p_1)$, we obtain a system with an intrinsic maximum throughput and such that the sum of initial markings $M(0, p_1)+M(0, p_2)$ is minimum.

Proof. By Lemma 1, the maximum throughput of the system equals the intrinsic throughput. So, we must prove that $M(0, p_1) + M(0, p_2)$ is minimum to achieve the throughput ρ.

By contradiction, let us suppose that there exists a marking M such that $M(0, p_1) + M(0, p_2) < 2 \cdot M_{min}(p_1)$ with a throughput equal to the intrinsic maximum value. According to lemma 2 page 6, there is a date $\tau^\star$ such that transitions t_i and t_j are both fired. Therefore, without loss of generality, we can consider that the marking M is such that $\tau^\star = 0$. Then, the initial marking of the corresponding WEG $\mathcal{G'}_{p_1,p_2}$ is well defined and verifies $M_0'(p_1) + M_0'(p_2) < v(p_1) + w(p_1) - 2 \cdot gcd_{p_1}$. So, by Lemma 5, the number of intermediary states is at most $N_i + N_j - 2$. But, by Lemmas 3 and 4, the number of intermediary states of $\mathcal{G'}_{p_1,p_2}$ is $N_i + N_j - 1$, the contradiction. □

Applying this last theorem to every couple of backward places of a SWTEG, we get the following corollary:

Corollary 1. *Let $\mathcal{G}$ be a unitary SWEGT such that*

$$\forall t_i \in T,\ \ \ell(t_i) = \rho \cdot Z_i\ ,\ \ \rho > 0.$$

Then setting $M(0, p) = M_{min}(p)$ for every place $p \in P$, we obtain a system with an intrinsic maximum throughput and such that $\sum_{p \in P} y_p \cdot M(0, p)$ is minimum.

We deduce an approximation algorithm for the general case:

Theorem 3. *Let $\mathcal{G}$ be a unitary SWTEG. Setting $M(0, p) = v(p) + w(p) - gcd_p\ \forall p \in P$ leads to a 2-approximation solution for the* Max Intrinsic Throughput *problem.*

Proof. Let $i^\star \in \{1, \ldots, n\}$ such that $\frac{Z_{i^\star}}{\ell(t_{i^\star})} = \min_{t_i \in T} \left(\frac{Z_i}{\ell(t_i)}\right)$. We set $\rho = \frac{Z_{i^\star}}{\ell(t_{i^\star})}$.

We first prove that the solution is live and that its throughput is equal to ρ. We show then that the criteria is at most twice from a minimum value.

1. Let us consider the periodic schedule $\sigma(t, k) = (k-1) \cdot \frac{Z_t}{\rho}$ for any $(t, k) \in T \times \mathbb{N}^\star$. Its troughput is ρ. Moreover, if we replace the duration of the firing sequences by values $\ell'(t_j) = \frac{Z_j}{\rho} \geq \ell(t_j)$, then the schedule σ is feasible for M by Lemma 1. So, we conclude that σ is feasible and that the system is live.
2. By (Marchetti & Munier-Kordon 2004), every live marking verify, for every couple (p, p') of backward places,

$$M(0, p) + M(0, p') \geq M_{min}(p).$$

So, $\frac{1}{2} \sum_{p \in P} y_p \cdot M_{min}(p)$ is a lower bound of the criteria and the ratio is obtained.

□

For the example pictured by Figure 6, we obtain the initial marking $M(0, p_1) = M(0, p_1') = 140$, $M(0, p_1) = M(0, p_1') = 100$, $M(0, p_1) = M(0, p_1') = 98$, $M(0, p_1) = M(0, p_1') = 126$.

Our purpose was initially to bound the capacity storage of places of Figure 2. We obtain then $C(p_1) = \frac{M(0,p_1)+M(0,p_1')}{x(p_1)} = \frac{280}{35} = 8$, $C(p_2) = \frac{M(0,p_2)+M(0,p_2')}{x(p_2)} = \frac{200}{10} = 20$, $C(p_3) = \frac{M(0,p_3)+M(0,p_3')}{x(p_3)} = \frac{196}{14} = 14$, and $C(p_4) = \frac{M(0,p_4)+M(0,p_4')}{x(p_4)} = \frac{252}{21} = 12$. The global surface on the chip is then equal to $8 \cdot y_{p_1} + 20 \cdot y_{p_2} + 14 \cdot y_{p_3} + 12 \cdot y_{p_4}$ and is less than twice from the optimum value. The restriction of M defines a live marking with throughput $\rho = \frac{35}{3}$.

Conclusions

We have developed in this paper a very simple approximation algorithm for Max Intrinsic Throughput problem. This new algorithm permits to give a first efficient solution to solve this industrial problem. Moreover, it will be useful to get good lower bounds to evaluate the performance of other algorithms.

References

Chrétienne, P. 1983. *Les réseaux de Petri temporisés*. Ph.D. Dissertation, Thèse d'état, Université P. et M. Curie.

Commoner, F.; Holt, A. W.; Even, S.; and Pnueli, A. 1971. Marked directed graphs. *J. Comput. Syst. Sci.* 5(5):511–523.

Di Febbraro, A.; Minciardi, R.; and Sacone, S. 1997. Deterministic timed event graphs for performance optimization of cyclic manufacturing processes. *IEEE Transactions on robotics and automation* 13(2).

Gaubert, S. 1990. An algebraic method for optimizing resources in timed event graphs. In *9th conference on Analysis and Optimization of Systems*, volume 144. Springer.

Giua, A.; Piccaluga, A.; and Seatzu, C. 2002. Firing rate optimization of cyclic timed event graphs. *Automatica* 38(1).

Govindarajan, R.; Gao, G. R.; and Desai, P. 2002. Minimizing memory requirements in rate-optimal schedule in regular dataflow networks. *Journal of VLSI Signal Processing* 31(3).

Hanen, C., and Munier, A. 1994. Cyclic scheduling on parallel processors: An overview. In Chrétienne, P.; Coffman, E. G.; Lenstra, J. K.; and Liu, Z., eds., *Scheduling theory and its applications*. J. Wiley and sons.

Karp, R. M., and Miller, R. E. 1966. Properties of a model for parallel computations: Determinacy, termination, queueing. *SIAM* 14(63):1390 – 1411.

Laftit, S.; Proth, J.-M.; and Xie, X. 1989. Optimization of invariant criteria for event graphs. *IEEE Transactions on Automatic Control* 37(5).

Lee, E. A., and Messerschmitt, D. G. 1987. Synchronous data flow. *IEEE Proceedings of the IEEE* 75(9).

Marchetti, O., and Munier-Kordon, A. 2004. A sufficient condition for the liveness of weighted event graphs. *Submitted, LIP6 Research Report, http://ftp.lip6.fr/lip6/reports/2005/lip6-2005-004.pdf*.

Marchetti, O., and Munier-Kordon, A. 2005. Minimizing places capacities of a weighted event graph. *Submitted, LIP6 Research Report, http://www-asim.lip6.fr/publications/2005/rp.lip6.2005.marchetti.pdf*.

Marchetti, O., and Munier-Kordon, A. 2006. Complexity results for bi-criteria cyclic scheduling problems. *Submitted, LIP6 Research Report, http://www-asim.lip6.fr/publications/resumes/2006/rp.2006.marchetti2/index.gb.html*.

Middendorf, M., and Timkovsky, V. G. 2002. On scheduling cycle shops: Classification, complexity and approximation. *Journal of Scheduling* 5(2):135–169.

Munier, A. 1993. Régime asymptotique optimal d'un graphe d'événements temporisé généralisé: application à un problème d'assemblage. *RAIRO-Automatique Productique Informatique Industrielle* 27(5):487–513.

Peterson, J.-L. 1981. *Petri Net Theory and the Modeling of Systems*. Prentice Hall PTR.

Sauer, N. 2003. Marking optimization of weighted marked graph. *Discrete Event Dynamic Systems* 13(3):245–262.

Stuijk, S.; Geilen, M.; and Basten, T. 2006. Exploring trade-offs in buffer requirements and throughput constraints for synchronous dataflow graphs. In *43rd Design Automation Conference, DAC 2006*. ACM Press.

Teruel, E.; Chrzastowski-Wachtel, P.; Colom, J. M.; and Silva, M. 1992. On weighted T-systems. In *Proocedings of the 13th Internationnal Conference on Application and Theory of Petri Nets 1992, Lecture Notes in Computer Science*, volume 616. Springer.