Linear Control of Live Marked Graphs

Philippe Darondeau, Xiaolan Xie

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Linear Control of Live Marked Graphs

Philippe Darondeau — Xiaolan Xie

N° 4251

July 2001

THÈMES 1 et 4

ISSN 0249-6399 ISRN INRIA/RR--4251--FR+ENG

Linear Control of Live Marked Graphs

Philippe Darondeau * , Xiaolan Xie †

Thèmes 1 et 4 — Réseaux et systèmes — Simulation et optimisation
de systèmes complexes
Projets S4 et Macsi

Rapport de recherche n° 4251 — July 2001 — 24 pages

Abstract: Given a linear constraint on the firing vectors of a live marked graph (bounded or unbounded) with uncontrollable / unobservable transitions, we apply linear programming techniques to compute the most liberal controller enforcing this constraint. In the special case of strongly connected live marked graphs, we compute further on the most liberal controller that keeps the marked graph live and enforces the constraint.

Key-words: Petri Nets, Marked Graphs, Firing Vectors, Linear Constraints, Supervision, Liveness, Polyhedra, Linear Programming

Research done within the Coordinated Research Action Mars of Inria

* irisa, campus de Beaulieu, F35042 Rennes Cedex (Philippe.Darondeau@irisa.fr)
† enim, île du Saulcy, F57045 Metz (Xiaolan.Xie@loria.fr)

Contrôle linéaire des graphes marqués vivants

Résumé : Etant donné une contrainte linéaire sur les vecteurs de tir d'un graphe marqué vivant (borné ou non) ayant des transitions incontrôlables / inobservables, nous utilisons les techniques de la programmation linéaire pour calculer le contrôleur le plus permissif mettant en oeuvre cette contrainte. Dans le cas de graphes marqués vivants fortement connexes, nous calculons en outre le contrôleur le plus permissif qui met en oeuvre la contrainte en préservant la vivacité.

Mots-clés : réseaux de Petri, graphes marqués, vecteurs de tir, contraintes linéaires, supervision, vivacité, polyèdres, programmation linéaire

1 Introduction

Enforcing linear constraints on reachable markings, and enforcing liveness, are yet two disjoint problems in Petri net supervisory control. Early work was developped in [2] [14] to enforce linear constraints on markings by auxiliary places, called *monitors*. Monitors achieve the most liberal control of safe marked graphs, but they fail in this respect when considering (bounded) Petri nets together with uncontrollable and unobservable transitions. Recent advances on monitors were made in two directions. One direction is to compute monitors by purely linear techniques so as to avoid constructing state graphs [7]: monitors obtained in this way do belong to the (integer) module generated by the places of the net (identified with rows of its incidence matrix). A second direction is to search for monitors in a larger space, namely the (integer) module of all regions of the state graph of the net (i.e. all implicit places of the net) [11]. Constructing state graphs is the price to pay for computing unrestricted monitors. A trade off between these two opposite directions has yet to be proposed. Pioneering work on liveness enforcement was done in [13], where an optimal control is produced in the form of a control net. Similar results were obtained recently in [4] in the extended context of nets with uncontrollable and unobservable transitions.

We shall in this paper propose a linear algebraic construction of the most permissive controller for a live marked graph, able to enforce a linear inequality on reachable markings or, more generally, on firing vectors, in the presence of uncontrollable and unobservable transitions. This, as far as we know, is a new contribution, since we address the case of unsafe and possibly unbounded live marked graphs, in which monitors fail to achieve the most permissive control. Moreover, linear inequalities on firing vectors allow to express a larger class of control objectives [5] [6]. As a matter of fact, many elements of our presentation are borrowed from the general framework defined by Li and Wonham, but our results diverge significantly from the results given in the last two references. On the one hand, marked graphs are a subclass of the Vector Discrete Event-Systems (VDES) considered there. On the other hand, Li and Wonham solve the control problem for VDES in the specific case when the uncontrollable projections of these systems are loop-free, and they use integer linear programming, while we solve the control problem for live marked graphs with arbitrary loops, and we use linear programming in the rational numbers. A disadvantage of the solutions we propose w.r.t. monitors is to lead to interpreted control, not to compiled control (making verification of the controlled system problematic). This is the price to pay if one wants to obtain the most permissive supervisory control in every case.

Another, more serious, limitation of the controllers we offer is that they fail to preserve liveness of the uncontrolled net. We do not know any solution to the problem of *jointly* enforcing linear constraints and preserving liveness for *unbounded* nets or marked graphs. We shall therefore limit our ambition to solve this problem for strongly connected live marked graphs, which are bounded. If we considered only linear constraints on reachable markings, the problem could be solved directly by iterative techniques on finite state systems, but the solution is not so direct when considering linear constraints on firing vectors.

The remaining sections are organized as follows. Section 2 examines sets of firing vectors of live marked graphs (they are polyhedra), and shows that they are stable under projection

on observable transitions. Sections 3 and 4 construct controllers for live marked graphs (using polyhedra, unimodularity and Farkas lemma). Section 5 compares our approach to the approach based on monitors. Sections 6 to 8 solve the question of liveness preserving control for strongly connected marked graphs (we fold polyhedra to polytopes and then extract geometric automata from the latter).

2 Live Marked Graphs

A *marked graph* is a Petri net in which all weights are 1 and every place has exactly one input transition and one output transition. Such nets may be seen as directed graphs, with transitions as nodes and places as arcs. A *marking* is a labeling of the arcs with non-negative integers. Let $G = (P, T, C, M_0)$ be a marked graph, where $C : P \times T \to \{-1, 0, 1\}$ is the *connectivity matrix* of the graph ($C[p,t]$ equals 1 if arc p leaves node t, -1 if arc p leads to node t, and 0 otherwise), and $M_0 : P \to \mathbb{N}$ is the *initial* marking (seen as a column-vector). For each transition $t \in T$, let $\mathbf{t} : T \to \mathbb{N}$ be the column-vector $\mathbf{t}(t) = 1$ and $\mathbf{t}(t') = 0$ for $t' \neq t$. Transition t may be *fired* at M_0 if and only if $M_0 + C\mathbf{t}$ is a non-negative vector (hence a marking of G). *Firing* t at M_0 is represented as $M_0[t\rangle M$, where $M = M_0 + C\mathbf{t}$. A marking M_n such that exists a firing sequence $M_0[t_1\rangle M_1[t_2\rangle \dots M_{n-1}[t_n\rangle M_n$ is said to be *reachable*. Abusing the notation, one writes also $M_0[t_1 \dots t_n\rangle M_n$. The sequence $\sigma = t_1 \dots t_n$ is called a *firing sequence* of G. The column-vector $\psi(\sigma) : T \to \mathbb{N}$ with entries $\psi(\sigma)(t)$ counting the occurrences of the respective transitions t in σ is called a *firing vector* of G. Further abusing the notation, let $M[\mathbf{x}\rangle M'$ be used to mean $\exists\sigma\ M[\sigma\rangle M'$ and $\psi(\sigma) = \mathbf{x}$.

Remarque 2.1 *For any firing vector* $\mathbf{x}$ *and transition* $t \in T$*, if* $M_0 + C(\mathbf{x} + \mathbf{t}) \geq 0$ *then* $\mathbf{x} + \mathbf{t}$ *is a firing vector.*

A marked graph is *live* if, for each firing vector $\mathbf{x}$ and for each transition $t \in T$, there exists $\sigma \in T^*$ such that $(M_0 + C\mathbf{x})[\sigma t\rangle M'$ for some marking M'. Thus, in a live marked graph, for each firing vector $\mathbf{x}$ and for each transition $t \in T$, there exists a firing vector $\mathbf{x}' \geq \mathbf{x}$ such that $(\mathbf{x}' - \mathbf{x})(t) \geq 1$ (comparison of, and operations on vectors are componentwise). Nevertheless, the latter condition may be satisfied also in a non-live marked graph. The sub-class of live marked graphs has a simple characterization: G is live if and only if the initial marking M_0 places at least one token on each directed circuit in the graph (see [9]). The dynamics of a live marked graph may be fully characterized in linear algebraic terms, as follows (Theo. 2 in [8]).

Theorem 2.2 *In a live marked graph* $G = (P, T, C, M_0)$*, a vector* $\mathbf{x} : T \to \mathbb{N}$ *is a firing vector if and only if* $M_0 + C\mathbf{x} \geq 0$ *– and in this case* $M_0\ [\mathbf{x}\rangle\ M_0 + C\mathbf{x}$*.*

Some consequences of this central theorem will be examined later on in the section. In the meantime, let us recall easy lemmas which belong to the folklore.

Lemma 2.3 (interpollation) *Given firing vectors* $\mathbf{x}$ *and* $\mathbf{x}'$ *of a live marked graph, let* M *and* M' *be the respective markings such that* $M_0[\mathbf{x}\rangle M$ *and* $M_0[\mathbf{x}'\rangle M'$. *If* $\mathbf{x} \leq \mathbf{x}'$, *then* $M[\mathbf{x}' - \mathbf{x}\rangle M'$.

Proof: Immediate from Theo. 2.2. ∎

Lemma 2.4 (least enabling vector) *In a live marked graph, given a reachable marking* M *and a transition* t, *there is a unique minimal vector* $\mathbf{x}$ *such that* $M[\mathbf{x}\rangle M'$ *for some marking* M' *enabling transition* t.

Proof: Let $\mathbf{x}(t') = 1$ for every transition (or node) t' located in between the extremities (excluding them) of a directed path of the graph, starting with an arc (or place) marked in M, taking no other arc marked in M, and leading to the target t; let $\mathbf{x}(t') = 0$ for all other transitions. ∎

Lemma 2.5 (meet semi-lattice) *The set of firing vectors of a live marked graph is closed under binary meet.*

Proof: Let $\mathbf{x}'$ and $\mathbf{x}''$ be two firing vectors of $G = (P, T, C, M_0)$. Suppose for the sake of contradiction that $\mathbf{x} = \mathbf{x}' \wedge \mathbf{x}''$ is not a firing vector of G. By Theo. 2.2, relation $M_0[\,p] + C[p, \cdot\,]\mathbf{x} \geq 0$ fails for some place $p \in P$, with associated row $C[\,p, \cdot\,]$ in matrix C. By definition of marked graphs, this row has two non-zero entries $C[\,p, t] = -1$ and $C[\,p, t'] = 1$. As $\mathbf{x}'$ and $\mathbf{x}''$ are firing vectors of G, by Theo. 2.2, $M_0[p] - \mathbf{x}'[t] + \mathbf{x}'[t'] \geq 0$ and $M_0[p] - \mathbf{x}''[t] + \mathbf{x}''[t'] \geq 0$. Therefore, *a fortiori*, $M_0[p] - \mathbf{x}[t] + \mathbf{x}'[t'] \geq 0$ and $M_0[\,p] - \mathbf{x}[t] + \mathbf{x}''[t'] \geq 0$, which entail altogether $M_0[p] - \mathbf{x}[t] + \mathbf{x}[t'] \geq 0$, a contradiction of the assumption. ∎

Lemma 2.6 (join semi-lattice) *The set of firing vectors of a live marked graph is closed under binary join.*

Proof: Let $\mathbf{x}'$ and $\mathbf{x}''$ be firing vectors of $G = (P, T, C, M_0)$. Put $\mathbf{x} = \mathbf{x}' \wedge \mathbf{x}''$, then vectors $\mathbf{x}' - \mathbf{x}$ and $\mathbf{x}'' - \mathbf{x}$ have disjoint subsets on non-zero entries. Therefore, $\mathbf{x}' \vee \mathbf{x}'' = \mathbf{x} + ((\mathbf{x}' - \mathbf{x}) \vee (\mathbf{x}'' - \mathbf{x})) = \mathbf{x} + (\mathbf{x}' - \mathbf{x}) + (\mathbf{x}'' - \mathbf{x})$. As each row $C[p, \cdot\,]$ of the connectivity matrix contains exactly two non-zero entries -1 and 1, the scalar products $C[p, \cdot\,](\mathbf{x}' - \mathbf{x})$ and $C[p, \cdot\,](\mathbf{x}'' - \mathbf{x})$ cannot be jointly negative. Therefore, $C[p, \cdot\,](\mathbf{x}' \vee \mathbf{x}'') \geq C[p, \cdot\,]\mathbf{x}' \wedge C[p, \cdot\,]\mathbf{x}''$. By Theo. 2.2, $M_0[p] + C[p, \cdot\,]\mathbf{x}' \geq 0$ and $M_0[p] + C[p, \cdot\,]\mathbf{x}'' \geq 0$, hence $M_0[p] + C[p, \cdot\,](\mathbf{x}' \vee \mathbf{x}'') \geq 0$ for every place p. By Theo. 2.2, $\mathbf{x}' \vee \mathbf{x}''$ is therefore a firing vector of G. ∎

Theorem 2.2 shows that the set of firing vectors of a live marked graph $G = (P, T, C, M_0)$ is the set of integral vectors of a *polyhedron*, defined by the system of linear inequalities $\mathbf{x} \geq 0$ and $-C\mathbf{x} \leq M_0$ (as $M_0 \geq 0$, the null vector is always a solution of the system). One may refine this observation. Recall that each row in the connectivity matrix C of a marked graph has exactly two non-zero entries -1 and 1. It was pointed out long ago that such matrices are *totally unimodular*, which means that the determinant of every square submatrix is always -1, 0 or 1. By Hoffman and Kruskal's theorem, it follows that the set

$\{\mathbf{x} \geq 0 \mid -C\mathbf{x} \leq M_0\}$ is an *integral polyhedron*, that is to say, this polyhedron is equal to the convex hull of the integral vectors it contains. We refer the reader to [12] for the general background of linear and integer programming. Classical results in this field may be used to show a nice property of sets of firing vectors of live marked graphs: this class of integral polyhedra is closed under projections. The rest of the section is devoted to establishing this fact, crucial for supervisory control of live marked graphs with unobservable transitions.

In the sequel, $G = (P, T, C, M_0)$ is a live marked graph whose set of transitions is the disjoint union $T = T' \cup T''$ of two subsets with respective numbers of elements $|T'| = m$ and $|T''| = n$. The latter subset contains *unobservable* transitions, to abstract from. We write $\mathbf{x} = \mathbf{x}'\mathbf{x}''$ to mean a non-negative rational column-vector $\mathbf{x}$ with the projections $\pi'\mathbf{x} = \mathbf{x}'$ and $\pi''\mathbf{x} = \mathbf{x}''$, respectively on T' and T''. We call $\pi'\mathbf{x} = \mathbf{x}'$ the *observable projection* of $\mathbf{x}$. Similarly, we write $C = C', C''$ to mean the block-decomposition of the connectivity matrix into submatrices $C' : P \times T' \to \{-1, 0, 1\}$ and $C'' : P \times T'' \to \{-1, 0, 1\}$. As usual, $\mathbb{Q}$ denotes the rational numbers, and $\mathbb{Q}_+$ denotes the non-negative rational numbers. In a first stage, we show that the observable projections of the firing vectors of a live marked graph are the integral points of a polyhedron.

Lemma 2.7 *The set* $\{\mathbf{x}' \in \mathbb{N}^m \mid (\exists \mathbf{x}'' \in \mathbb{N}^n) \; -C\mathbf{x}'\mathbf{x}'' \leq M_0\}$ *is equal to the set* $\{\mathbf{x}' \in \mathbb{N}^m \mid (\exists \mathbf{x}'' \in {\mathbb{Q}_+}^n) \; -C\mathbf{x}'\mathbf{x}'' \leq M_0\}$

Proof: The inclusion of the former set in the latter is trivial. In order to prove the converse inclusion, it suffices to show that whenever $-C\mathbf{x}'\mathbf{q}'' \leq M_0$ for a non-negative rational vector $\mathbf{q}''$, then $-C\mathbf{x}'\mathbf{x}'' \leq M_0$ for $\mathbf{x}'' = \lceil \mathbf{q}'' \rceil$. For this purpose, consider any inequality $-C[p, \cdot\,]\,\mathbf{x} \leq M_0[p]$ in system $-C\mathbf{x} \leq M_0$. Because row $C[p, \cdot\,]$ has exactly two non-zero entries -1 and 1, $-C''[p, \cdot\,]\,(\mathbf{x}'' - \mathbf{q}'') < 1$. As $-C[p, \cdot\,]\,\mathbf{x}'\mathbf{q}'' \leq M_0[p]$, it comes that $-C[p, \cdot\,]\,\mathbf{x}'\mathbf{x}'' \leq M_0[p]$ since both numbers are integer and the left one is strictly smaller than $M_0[p] + 1$. ∎

Proposition 2.8 *The observable projections of the firing vectors of a live marked graph are the integral points of a polyhedron.*

Proof: Given $\mathbf{x}' \in \mathbb{N}^m$, by the above lemma, $\mathbf{x}'$ is the observable projection of a firing vector if and only if $(\exists \mathbf{q}'' \in {\mathbb{Q}_+}^n) \; -C''\mathbf{q}'' \leq M_0 + C'\mathbf{x}'$. By Farkas lemma, $-C''\mathbf{q}'' \leq M_0 + C'\mathbf{x}'$ has a non-negative solution if and only if, for every rational row-vector $\mathbf{y}$ with dimension equal to the number of places $|P|$:

$$\mathbf{y}(-C'') \geq 0 \text{ and } \mathbf{y} \geq 0 \;\Rightarrow\; \mathbf{y}(M_0 + C'\mathbf{x}') \geq 0$$

Now the conditions $\mathbf{y}(-C'') \geq 0$ and $\mathbf{y} \geq 0$ define a convex cone, and the linear homogeneous inequality $\mathbf{y}(M_0 + C'\mathbf{x}') \geq 0$ holds for all vectors $\mathbf{y}$ in this cone if and only if holds for its extremal rays $\mathbf{y}_1, \ldots, \mathbf{y}_k$. Therefore, vector $\mathbf{x}'$ is the observable projection of a firing vector if and only if it belongs to the polyhedron defined by the linear inequalities $-(\mathbf{y}_i\, C')\,\mathbf{x}' \leq \mathbf{y}_i\, M_0 \;\; (i = 1 \ldots k)$. ∎

We will now show that the integral points of the above polyhedron are the firing vectors of a live marked graph. Therefore, we must show that rational vectors $\mathbf{y}_1, \ldots, \mathbf{y}_k$ representing the extremal rays of the convex cone $\mathbf{y} \geq 0, \mathbf{y}\, C'' \leq 0$ may be chosen such that each linear combination $\mathbf{y}_i\, C'$ of rows of C' contains exactly two non-zero entries -1 and 1. Clearly, the extremal rays of a convex cone may be represented as integral vectors, so let us focus on integral solutions of system $\mathbf{y} \geq 0, \mathbf{y}\, C'' \leq 0$. Exploiting the strong fact that C'' is a sub-matrix of a connectivity matrix, we prove that each integral solution $\mathbf{y}$ is a non-negative linear combination of smaller solutions with all entries in $\{0, 1\}$.

In next lemma, we let P_{io}, P_{oo}, P_{oi}, and P_{ii} denote the subsets of arcs (or places) of the marked graph G that lead, respectively, from unobservable nodes in T'' to observable nodes in T', from observable nodes to observable nodes, from observable nodes in T' to unobservable nodes in T'', and from unobservable nodes to unobservable nodes.

Lemma 2.9 *Each non-negative integral solution* $\mathbf{y}$ *of* $\mathbf{y}\, C'' \leq 0$ *expresses as a sum of* $\{0,1\}$*-solutions* $\mathbf{y}'$ *satisfying one of four mutually exclusive conditions:*
i) $(\exists p' \in P_{oo})\ \ \mathbf{y}'[p] = 1$ iff $p = p'$,
ii) $(\exists P' \subseteq P_{ii})\ \ \mathbf{y}'[p] = 1$ iff $p \in P'$,
iii) $(\exists P' \subseteq P_{ii})\ (\exists p' \in P_{oi})\ \ \mathbf{y}'[p] = 1$ iff $(p \in P'$ or $p = p')$,
iv) $(\exists P' \subseteq P_{ii})\ (\exists p' \in P_{oi})\ (\exists p'' \in P_{io})\ \ \mathbf{y}'[p] = 1$ iff $(p \in P'$ or $p \in \{p', p''\})$.

Proof: Consider first the simple case when every row $C''[p, \cdot\,]$ with a positive coefficient $\mathbf{y}[p] > 0$ has no entry equal to 1. Then either all entries of $C''[p, \cdot\,]$ are zeroes and $p \in P_{oo}$, or $C''[p, t] = -1$ for some $t \in T''$ and $p \in P_{oi}$. Therefore, $\mathbf{y}$ is a sum of smaller solutions of type (i) or (iii). Consider now the case when exists a place p_1 such that $\mathbf{y}[p_1] \geq 1$ while $C''[p_1, t_1] = 1$ for some transition t_1. As $\mathbf{y}\, C'' \leq 0$, there must exist a place p_2 such that $\mathbf{y}[p_2] \geq 1$ and $C''[p_2, t_1] = -1$. If $C''[p_2, t_2] = 1$ for some transition t_2, then, there must exist a place p_3 such that $\mathbf{y}[p_3] \geq 1$ and $C''[p_3, t_2] = -1$ (where possibly $p_3 = p_1$). Continuing the iteration, and producing at each step a place p_{k+1} such that $\mathbf{y}[p_{k+1}] \geq 1$ and $C''[p_{k+1}, t_k] = -1$, one meets sooner or later case (a) or case (b):
a) $p_{k+1} = p_h$ with $h < k$,
b) $C''[p_{k+1}, \cdot\,]$ has no entry equal to 1.
In case (a), let $\mathbf{y}'$ be the $\{0,1\}$-vector defined with $\mathbf{y}'[p] = 1$ iff $p \in \{p_h, \ldots, p_k\}$. Then $\mathbf{y}'\, C'' = 0$, hence $\mathbf{y}'$ is a solution of type (ii), and one may proceed to decompose $(\mathbf{y} - \mathbf{y}')$. In case (b), one distinguishes two subcases. The easy case is when $C''[p_1, \cdot\,]$ has no entry equal to -1. In this case, let $\mathbf{y}'$ be the $\{0,1\}$-vector defined with $\mathbf{y}'[p] = 1$ iff $p \in \{p_1, \ldots, p_{k+1}\}$. Then $\mathbf{y}'\, C'' = 0$, hence $\mathbf{y}'$ is a solution of type (iv) – with $p' = p_{k+1}$, $P' = \{p_2, \ldots, p_k\}$, and $p'' = p_1$ – and one may proceed to decompose $(\mathbf{y} - \mathbf{y}')$. The uneasy case is when $C''[p_1, t_0] = -1$ for some transition t_0. We distinguish further two subcases. The first case is when $(\mathbf{y}\, C'')[t_0] \leq -1$. Then, let $\mathbf{y}'$ be the $\{0,1\}$-vector defined with $\mathbf{y}'[p] = 1$ iff $p \in \{p_1, \ldots, p_{k+1}\}$. As $\mathbf{y}'\, C''$ has a unique non-zero entry $(\mathbf{y}'\, C'')[t_0] = -1$, $\mathbf{y}'$ is a solution of type (iii) – with $P' = \{p_1, \ldots, p_k\}$ and $p' = p_{k+1}$. Now, $(\mathbf{y} - \mathbf{y}')C'' \leq 0$, and one may proceed to decompose $(\mathbf{y} - \mathbf{y}')$. The second case is when $(\mathbf{y}\, C'')[t_0] = 0$. Then, there must exist a place p_0 such that $\mathbf{y}[p_0] \geq 1$ and $C''[p_0, t_0] = 1$. If $p_0 = p_h \in \{p_1, \ldots, p_{k+1}\}$, let $\mathbf{y}'$

be the $\{0,1\}$-vector defined with $\mathbf{y}'[p] = 1$ iff $p \in \{p_1, \ldots, p_h\}$. Then $\mathbf{y}' C'' = 0$, hence $\mathbf{y}'$ is a solution of type (ii), and one may proceed to decompose $(\mathbf{y} - \mathbf{y}')$. If on the contrary $p_0 \notin \{p_1, \ldots, p_{k+1}\}$, one is brought back to case (b) with the new place p_0 playing the role of the former place p_1, hence one iterates the reasoning. ∎

Proposition 2.10 *Observable projections of firing vectors of a live marked graph are the integral points of a polyhedron $\mathbf{x}' \geq 0, -D\mathbf{x}' \leq M'$ where each row of D has two non-zero entries -1 and 1, and M' is a vector of non-negative integers.*

Proof: By the above lemma, the extremal rays of the cone $\mathbf{y} \geq 0, \mathbf{y} C'' \leq 0$ may be represented by $\{0,1\}$-vectors $\mathbf{y}_i$ of types (i) to (iv). Seeing that $C'[p, \cdot]$ is an all-zero vector whenever $p \in P_{ii}$, it follows that each non-trivial inequality $-(\mathbf{y}_i C')\,\mathbf{x}' \leq \mathbf{y}_i M_0$ boils down to one of the forms:
i) $-C'[p, \cdot]\,\mathbf{x}' \leq M_0[p]$ where $p \in P_{oo}$,
thus $-C'[p, \cdot]$ has two non-zero entries -1 and 1,
iii) $-C'[p', \cdot]\,\mathbf{x}' \leq M_0[p']$ where $p' \in P_{oi}$,
thus $C'[p', \cdot]$ is a non-negative vector and the inequality is trivial,
iv) $-C'[p', \cdot]\,\mathbf{x}' - C'[p'', \cdot]\,\mathbf{x}' \leq M_0[p'] + M_0[p'']$ where $p' \in P_{oi}$ and $p'' \in P_{io}$, thus $C'[p', \cdot]$ has a unique non-zero entry, equal to 1, and $C'[p'', \cdot]$ has a unique non-zero entry, equal to -1, and $(-C'[p', \cdot]) + (-C'[p'', \cdot])$ either is an all-zero vector or has two non-zero entries -1 and 1. ∎

Theorem 2.11 *The observable projections of the firing vectors of a live marked graph are the firing vectors of a live marked graph.*

Proof: Let G' be the marked graph with the connectivity matrix D and the initial marking M'. Denote by $FV(G)$ and $FV(G')$ the respective sets of firing vectors of G and G'. By construction, $-D\mathbf{x}' \leq M'$ for every $\mathbf{x}' \in FV(G')$, hence $FV(G') \subseteq \{\pi'\mathbf{x} \,|\, \mathbf{x} \in FV(G)\}$. We claim that $FV(G') = \{\pi'\mathbf{x} \,|\, \mathbf{x} \in FV(G)\}$. As the null vector belongs to $FV(G')$, the right-to-left inclusion may be established by an induction on the non-negative integer vectors $\mathbf{x}'$ such that $-D\mathbf{x}' \leq M'$. In view of remark 2.1, the induction step amounts to showing that whenever $\mathbf{x}' \neq 0$, there exists some transition $t' \in T'$ such that $\mathbf{x}' \geq \mathbf{t}'$ and $-D\mathbf{y}' \leq M'$ for $\mathbf{y}' = \mathbf{x}' - \mathbf{t}'$. Now, $\mathbf{x}' = \pi'\mathbf{x}$ for some $\mathbf{x} \in FV(G)$, and $\pi'\mathbf{x} \neq 0$ entails that $\{\mathbf{y} \in FV(G) \,|\, \pi'\mathbf{y} < \pi'\mathbf{x}\}$ is non-empty. Choose vector $\mathbf{y}$ maximal in this set, then necessarily $\pi'\mathbf{x} - \pi'\mathbf{y} = \mathbf{t}'$ for some $t' \in T'$, and $-D\mathbf{y}' \leq M'$ for $\mathbf{y}' = \pi'\mathbf{y}$ since $\mathbf{y}'$ is the observable projection of a firing vector of G.

It remains to show that G' is live. As G is live and any reachable marking of G' is equal to $M' + D(\pi'\mathbf{y})$ for some $\mathbf{y} \in FV(G)$, it suffices to show that, if $\mathbf{y} \leq \mathbf{x} \in FV(G)$, then $(M' + D(\pi'\mathbf{y}))\ [\sigma\rangle\ (M' + D(\pi'\mathbf{x}))$ in G' for some sequence σ with firing count equal to $\pi'(\mathbf{x} - \mathbf{y})$. By induction, if suffices to consider the case where $\mathbf{x} - \mathbf{y} = \mathbf{t}$ for $t \in T$. As $\pi'\mathbf{x}$ is a firing vector of G', $M' + D(\pi'\mathbf{x})$ is a well-defined marking of G'. Thus, it suffices to put down $\sigma = t$ if t is observable ($t \in T'$), and to let σ be the empty sequence otherwise. ∎

Corollary 2.12 *The observable projections of the firing sequences of a live marked graph are the firing sequences of a live marked graph.*

Proof: It suffices to show that, given two firing vectors $\mathbf{x'}_1$ and $\mathbf{x'}_2$ of G' such that $\mathbf{x'}_1 \leq \mathbf{x'}_2$ and $\mathbf{x'}_2 - \mathbf{x'}_1 = \mathbf{t'}$ for some observable transition t', there exists firing vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ of G such that $\mathbf{x'}_1 = \pi'\mathbf{x}_1$, $\mathbf{x'}_2 = \pi'\mathbf{x}_2$, and $\mathbf{x}_1 \leq \mathbf{x}_2$. Choose any firing vectors $\mathbf{x}_1$ and $\mathbf{y}$ of G with the respective observable projections $\mathbf{x'}_1$ and $\mathbf{x'}_2$. By lemma 2.6, $\mathbf{x}_2 = \mathbf{x}_1 \vee \mathbf{y}$ is a firing vector of G, and $\pi'\mathbf{x}_2 = \mathbf{x'}_2$. ∎

3 The supervision problem for live marked graphs

Given a live marked graph $G = (P, T, C, M_0)$, or *plant*, let a subset of *legal* firing vectors be defined by linear inequality $A\mathbf{x} \leq b$. Let the set of transitions $T = T_c \cup T_u$ be partitioned into *controllable* transitions (T_c) and *uncontrollable* transitions (T_u). Finally, let $T_u = T_f \cup T_i$ be partitioned into *free* transitions (T_f) and *invisible* or *unobservable* transitions (T_i). Thus, $T_c \cup T_f$ is the subset of the *observable* transitions. The problem is to construct a process, the *supervisor*, able to confine the executions of plant G within the constraint $A\mathbf{x} \leq b$ by either granting or denying a permission for each occurrence of a controllable action of the plant, based on the awareness of all occurrences of observable actions.

This problem is a close variation of the problem usually considered in the literature of Petri net supervision, where legal *markings* are specified by a linear constraint $A\mathbf{M} \leq b$ ([2] [14] [3] [7]). As a matter of fact, a linear constraint on markings may be formulated alternatively as a linear constraint on firing vectors: the reachable markings are the images of the firing vectors under the linear transformation $M = M_0 + C\mathbf{x}$, hence constraint $A\mathbf{M} \leq b$ may also be expressed in the form $(AC)\,\mathbf{x} \leq (b - AM_0)$. What is gained by considering linear constraints on firing vectors is more power for expressing control objectives: one can discriminate between legal and illegal firing vectors that may lead to the same marking. Linear constraints on firing vectors stay however weaker than rational language constraints dealt with in [10]. We borrow for the rest the general framework proposed by Ramadge and Wonham for studying the supervision of discrete event systems with uncontrollable and unobservable events.

Our main objective in this section is to show that when the plant is a live marked graph, the sole information that a control process should maintain on what happened in the plant is the observable projection of the current firing vector. We will also determine which occurrences of the controllable actions should be granted or denied permission according to the most permissive policy. Computational aspects of control will be dealt with in next section.

Let notations as follows. The subset of *observable* transitions is denoted T_o, thus $T_o = T \setminus T_i = T_c \cup T_f$ (controllable transitions are observable, free transitions are observable but not controllable). For $s \in \{c, f, i, o\}$, let $\phi_s : T^* \to T_s^*$ be the morphism of free monoids defined with $\phi_s(t) = t$ for $t \in T_s$ and $\phi_s(t) = \varepsilon$ for $t \in T \setminus T_s$ – where ε is the empty word. Thus ϕ_s erases letters not in T_s. For $s \in \{c, f, i, o\}$, let $\psi_s : T_s^* \to (T_s \to \mathbb{N})$ denote the map that sends each word to the vector with entries counting occurrences of each letter in the word, and let $\psi : T^* \to (T \to \mathbb{N})$ be defined similarly for arbitrary words in T^*. Finally, for $s \in \{c, f, i, o\}$, let $\pi_s : (T \to \mathbb{N}) \to (T_s \to \mathbb{N})$ denote the projection that sends

each T-vector to its induced restriction on the entries in T_s. Thus, $\pi_s \circ \psi(w) = \psi_s \circ \phi_s(w)$ for all $w \in T^*$ and $s \in \{c, f, i, o\}$. In the sequel, *plant* refers to a fixed live marked graph $G = (P, T, C, M_0)$. $FS(G)$ denotes the set of all firing sequences of G, and $FV(G)$ denotes the set of all its firing vectors.

Proposition 3.1 *Let $\sigma', \sigma'' \in T^*$ be firing sequences of G such that $\psi_o \circ \phi_o(\sigma') = \psi_o \circ \phi_o(\sigma'')$, then there exists a firing sequence σ such that $\phi_o(\sigma) = \phi_o(\sigma')$ and $\psi_i \circ \phi_i(\sigma) = \psi_i \circ \phi_i(\sigma'')$.*

Proof: Let $\sigma' = w_1 t_1 \dots w_n t_n w_{n+1}$, where the w_j are sequences of unobservable transitions ($w_j \in T_i^*$ for $j = 1 \dots n+1$) and the t_j are observable transitions ($t_j \in T_o$ for $j = 1 \dots n$). One may assume $w_{n+1} = \varepsilon$ without loss of generality, as the statement to prove depends only upon $\phi_o(\sigma') = t_1 \dots t_n$. For the same reason, if we let $M_{j-1} = M_0 + C\psi(w_1 t_1 \dots t_{j-1})$ for $j > 1$, one may assume without loss of generality that, for all $j \in \{1, \dots, n\}$, $\psi(w_j)$ is the least vector, given by lemma 2.4, such that $M_{j-1}[\mathbf{x}_j\rangle M'_j$ for some marking M'_j enabling t_j. Actually, if the assumption holds for all $j < k$ but it does not hold for $j = k$, let w'_k be any sequence of transitions such that $\psi(w'_k) = \mathbf{x}_k$ (the least vector that must be fired from M_{k-1} in order to enable t_k). Thus, $M_{k-1}[w'_k\rangle M'_k[t_k\rangle$ for some M'_k, and at the same time, $M_{k-1}[w_k t_k\rangle M_k$, with $\psi(w'_k) = \mathbf{x}_k < \psi(w_k)$. It follows that $\psi(w'_k t_k) < \psi(w_k t_k)$, hence, by lemma 2.3, there exists a transition sequence w''_k with the commutative image $\psi(w''_k) = \psi(w_k t_k) - \psi(w'_k t_k) = \psi(w_k) - \psi(w'_k)$ such that $M_{k-1}[w'_k\, t_k\, w''_k\rangle M_k$. Therefore, the inductive assumption now holds up to $j = k$ in the modified sequence $w_1 t_1 \dots t_{k-1} w'_k\, t_k\, w''_k\, w_{k+1} t_{k+1} \dots$, whose observable trace $t_1 \dots t_n$ is still equal to $\phi_o(\sigma')$.

Let us proceed to the proof under these assumptions. As $\psi(w_1)$ is the least vector that must be fired from M_0 in order to enable t_1, and t_1 occurs in σ'' (this follows from the relation $\psi_o \circ \phi_o(\sigma') = \psi_o \circ \phi_o(\sigma'')$), one has necessarily $\psi(w_1 t_1) \leq \psi(\sigma'')$. By lemma 2.3, there exists therefore a transition sequence σ''_1 such that $M_1[\sigma''_1\rangle$ and $\psi(\sigma''_1) = \psi(\sigma'') - \psi(w_1 t_1)$. Thus, if we denote by σ'_1 the sequence such that $\sigma' = w_1\, t_1\, \sigma'_1$, we obtain $\psi_o \circ \phi_o(\sigma'_1) = \psi_o \circ \phi_o(\sigma') - \psi_o \circ \phi_o(w_1 t_1) = \psi_o \circ \phi_o(\sigma'') - \psi_o \circ \phi_o(w_1 t_1) = \pi_o \circ \psi(\sigma'') - \pi_o \circ \psi(w_1 t_1) = \pi_o(\psi(\sigma'') - \psi(w_1 t_1)) = \pi_o \circ \psi(\sigma''_1) = \psi_o \circ \phi_o(\sigma''_1)$. Therefore, if we consider now transition sequences σ'_1 and σ''_1, that can be fired at M_1, we can iterate the reasoning. At last stage in this iterative process, one obtains a sequence σ''_n such that $M_n[\sigma''_n\rangle$ and $\psi(\sigma''_n) = \psi(\sigma'') - \psi(w_1 t_1) - \dots - \psi(w_n t_n) = \psi(\sigma'') - \psi(\sigma')$. Thus, if we define $\sigma = w_1\, t_1 \dots w_n\, t_n\, \sigma''_n$, one finally derives $\psi_i \circ \phi_i(\sigma) = \pi_i \circ \psi(\sigma) = \pi_i(\psi(\sigma') + \psi(\sigma''_n)) = \pi_i(\psi(\sigma') + (\psi(\sigma'') - \psi(\sigma'))) = \pi_i \circ \psi(\sigma'') = \psi_i \circ \phi_i(\sigma'')$. ∎

Corollary 3.2 *For any sequence $\sigma \in FS(G)$, $\{\psi(\sigma') \mid \psi_o \circ \phi_o(\sigma') = \psi_o \circ \phi_o(\sigma)\} = \{\psi(\sigma') \mid \phi_o(\sigma') = \phi_o(\sigma)\}$, where σ' ranges over $FS(G)$.*

Proof: The inclusion of the latter set in the former is immediate. In order to show direct inclusion, suppose $\psi_o \circ \phi_o(\sigma') = \psi_o \circ \phi_o(\sigma)$. Then, by Prop. 3.1, there exists a sequence $\sigma'' \in FS(G)$ such that $\phi_o(\sigma'') = \phi_o(\sigma)$ and $\psi_i \circ \phi_i(\sigma'') = \psi_i \circ \phi_i(\sigma')$. From the first equality, $\psi_o \circ \phi_o(\sigma') = \psi_o \circ \phi_o(\sigma'')$, hence it follows from the second equality that $\psi(\sigma') = \psi(\sigma'')$. ∎

The above corollary shows that there is no need, for controlling a plant, to maintain an ordered history of occurrences of observable transitions: ordered histories $\phi_o(\sigma)$ and unordered histories $\psi_o \circ \phi_o(\sigma)$ determine the same predicates for identifying the firing vector $\psi(\sigma)$ in $FV(G)$.

Let us now turn to the topic of control policies. We are interested in the most permissive control policy, meaning that a controllable transition is never denied permission unless this would open a way in which the plant could subsequently violate the constraint $A\mathbf{x} \leq b$ against any defense, i.e. whatever could be the control policy applied after granting permission. Some terminology is introduced now in order to help describing this most permissive control policy.

Definition 3.3 *A firing vector* $\mathbf{x} \in FV(G)$ *is* forbidden *if* $A\mathbf{x} > b$, *it is* risky *if* $\mathbf{x} \leq \mathbf{x}'$ *and* $\pi_c(\mathbf{x}) = \pi_c(\mathbf{x}')$ *for some forbidden vector* $\mathbf{x}'$, *and it is* suspect *if* $\pi_o(\mathbf{x}) = \pi_o(\mathbf{x}')$ *for some risky vector* $\mathbf{x}'$. *Two vectors such that* $\pi_o(\mathbf{x}) = \pi_o(\mathbf{x}')$ *are* observationally equivalent, *noted* $\mathbf{x} \simeq_o \mathbf{x}'$.

Note that *forbidden* $\subseteq$ *risky* $\subseteq$ *suspect*. In view of lemma 2.3, a control policy that does not prevent the plant from firing all risky vectors cannot be enforce the control objective $A\mathbf{x} \leq b$. Actually, if the plant was allowed to perform process $M_0[\mathbf{x}\rangle M$ even though $\mathbf{x} \leq \mathbf{x}'$ and $\pi_c(\mathbf{x}) = \pi_c(\mathbf{x}')$ for some forbidden vector $\mathbf{x}'$, the control policy could not obstruct to $M[\mathbf{x}' - \mathbf{x}\rangle$, since all transitions in this process are uncontrollable. Next proposition shows that in the specific case of live marked graphs, an effective control policy should also prevent the plant from firing all suspect vectors.

Proposition 3.4 *Any deterministic policy for granting or denying permission to controllable transitions of the plant, based on the observation* $\phi_o(\sigma)$ *of the current transition sequence* σ *fired in the plant, either fails to enforce the control objective* $A\mathbf{x} \leq b$, *or results in preventing the plant from firing all suspect vectors.*

Proof: Consider a transition sequence σ of the plant such that $\psi(\sigma)$ is suspect. Assume that $M_0[\sigma\rangle$ is not obstructed by the control policy. Then, by definition, $\psi(\sigma) \simeq_o \psi(\sigma')$ for some transition sequence σ' such that $\psi(\sigma')$ is a risky vector. By Prop. 3.1, there exists yet another transition sequence σ'' of the plant such that $\phi_o(\sigma'') = \phi_o(\sigma)$ and $\psi(\sigma'') = \psi(\sigma')$. Therefore, $\psi(\sigma'')$ is also a risky vector. Let $\sigma'' = w_1 t_1 \ldots w_n t_n w_{n+1}$ where the w_j are sequences of unobservable transitions (for $j = 1 \ldots n+1$) and the t_j are observable transitions (for $j = 1 \ldots n$). As the control policy is deterministic, and whenever t_j is controllable, it has been granted permission to fire in run σ based on the observation $t_1 \ldots t_{j-1}$, the same must be true in run σ''. Therefore, the control policy does not prevent the plant from firing the risky vector $\psi(\sigma'')$, and it cannot enforce the control objective $A\mathbf{x} \leq b$ ∎

Proposition 3.5 *Let* $\mathbf{x}, \mathbf{x}'$ *be two firing vectors and* t *an uncontrollable transition such that* $\mathbf{x} \leq \mathbf{x}'$ *and* $\mathbf{x}' - \mathbf{x} = \mathbf{t}$ *(we recall that* $\mathbf{t}(t) = 1$ *and* $\mathbf{t}(t') = 0$ *for* $t' \neq t$*). If* $\mathbf{x}'$ *is a suspect vector, then* $\mathbf{x}$ *is a suspect vector.*

Proof: By definition, $\mathbf{x}' \simeq_o \mathbf{x}''$ for some risky vector $\mathbf{x}''$. Transition t may be either observable or unobservable. If t is unobservable ($t \in T_i$), then clearly, $\pi_o(x) = \pi_o(x') = \pi_o(x'')$, showing that $\mathbf{x}$ is suspect. Consider now the case when t is observable ($t \in T_u \cap T_o = T_f$). Select a firing sequence σ of G such that $\psi(\sigma) = \mathbf{x}$. As $\psi(\sigma t) = \mathbf{x}'$, σt is a firing sequence of G. As $\mathbf{x}' \simeq_o \mathbf{x}''$, by Prop. 3.1, there exists a firing sequence σ'' of G such that $\phi_o(\sigma'') = \phi_o(\sigma t)$ and $\psi(\sigma'') = \mathbf{x}''$. Let $\sigma'' = \alpha\, t\, \beta$ where β is a sequence of unobservable transitions. Because $t\,\beta$ is a sequence of uncontrollable transitions and $\mathbf{x}''$ is a risky vector, $\psi(\alpha)$ is a risky vector. Now, $\phi_o(\sigma'') = \phi_o(\alpha\, t)$ and $\phi_o(\sigma'') = \phi_o(\sigma\, t)$ entail $\phi_o(\alpha) = \phi_o(\sigma)$. Therefore, $\pi_o \circ \psi(\alpha) = \psi_o \circ \phi_o(\alpha) = \psi_o \circ \phi_o(\sigma) = \pi_o \circ \psi(\sigma)$, showing that $\mathbf{x} = \psi(\sigma)$ is a suspect vector. ∎

We are now ready to describe the most permissive control policy: *when the sequence σ has been fired in the plant, grant the permission to fire controllable action t if and only if $\psi(\sigma\, t)$ is not a suspect vector.* By Prop. 3.5, this control policy is correct (*forbidden* $\subseteq$ *suspect*). By Prop. 3.4, any correct control policy must prevent the plant from firing all suspect vectors, hence this control policy is the most permissive. What remains to be shown is that it is feasible, based on the observation $\phi_o(\sigma)$ of the sequence σ fired in the plant. As the set of the suspect vectors is closed under observational equivalence $\simeq_o$, one can actually determine from the observed sequence $\phi_o(\sigma)$ whether $\psi(\sigma\, t)$ is suspect. Even better, one can determine this from the data $\psi_o \circ \phi_o(\sigma) = \pi_o \circ \psi(\sigma)$. Therefore, the most permissive control policy may be realized by a process whose parameter is the observable projection of the vector yet fired in the plant (confirming the remark after corollary 3.2). Prop. 2.8 determines the range of this parameter, and corollary 2.12 explains its interpretation.

4 Computing linear control of live marked graphs

We show in this section that observable projections of suspect vectors are the integral points of a polyhedron. An effective computation of this polyhedron defines a linear control of plant $G = (P, T, C, M_0)$, enforcing linear constraint $A\mathbf{x} \leq b$ in the most permissive way.

Adapting notations from section 2, let $\mathbf{x} = \mathbf{x}_c\, \mathbf{x}_f\, \mathbf{x}_i = \mathbf{x}_c\, \mathbf{x}_u = \mathbf{x}_o\, \mathbf{x}_i$ denote a vector $\mathbf{x} : T \to \mathbb{N}$ with the respective projections $\pi_s(\mathbf{x}) = \mathbf{x}_s$ for $s \in \{c, f, i, o, u\}$. Similarly, let $\mathbf{q} = \mathbf{q}_c\, \mathbf{q}_f\, \mathbf{q}_i$ denote a rational vector $\mathbf{q} : T \to \mathbb{Q}$ with the respective projections $\pi_s(\mathbf{q}) = \mathbf{q}_s$. Finally, let $C_s : P \times T_s \to \{-1, 0, 1\}$ be the induced restriction of connectivity matrix C on subset of columns T_s.

It is easily seen that a firing vector $\mathbf{x}$ is suspect if and only if there exists a forbidden vector $\mathbf{x}'$ such that $\mathbf{x}_c = \mathbf{x}'_c$ and $\mathbf{x}_f \leq \mathbf{x}'_f$. Actually, in this case, $\mathbf{x}$ is observationaly equivalent to the risky vector $\mathbf{x} \wedge \mathbf{x}'$ given by lemma 2.5; conversely, if $\mathbf{x} \simeq_o \mathbf{x}''$ and $\mathbf{x}''$ is a risky vector, then there exists a forbidden vector $\mathbf{x}'$ such that $\mathbf{x}''_c = \mathbf{x}'_c$ and $\mathbf{x}''_u \leq \mathbf{x}'_u$, entailing $\mathbf{x}_c = \mathbf{x}'_c$ and $\mathbf{x}_f \leq \mathbf{x}'_f$. We shall now propose a linear characterization of suspect vectors, relying on total unimodularity of matrix C.

Lemma 4.1 *The firing vector $\mathbf{x} = \mathbf{x}_c\, \mathbf{x}_f\, \mathbf{x}_i$ is suspect if and only if there exist non-negative rational vectors $\mathbf{q}_f$ and $\mathbf{q}_i$ such that $\mathbf{x}_f \leq \mathbf{q}_f$, $-C(\mathbf{x}_c\, \mathbf{q}_f\, \mathbf{q}_i) \leq M_0$, and $A(\mathbf{x}_c\, \mathbf{q}_f\, \mathbf{q}_i) \geq b + 1$.*

Proof: In view of the observation made above, the specified conditions are necessary. Let us show that they are also sufficient. For this purpose, consider the linear form $A_u\,\mathbf{q}_u$ in the rational vector variable $\mathbf{q}_u = \mathbf{q}_f\,\mathbf{q}_i$, subject to linear constraints $-\mathbf{q}_f \leq -\mathbf{x}_f$ and $-C_u\,\mathbf{q}_u \leq M_0 + C_c\,\mathbf{x}_c$. The two constraints may be gathered into a single system $D\,\mathbf{q}_u \leq \mathbf{e}$, where matrix D has block decomposition:

$$\begin{pmatrix} -I_f & 0_i \\ -C_f & -C_i \end{pmatrix}$$

and $\mathbf{e}$ is the obvious vector of integers. Now, matrix C is totally unimodular, hence the opposite matrix $-C$ and its sub-matrix $-C_u = (-C_f, -C_i)$ are totally unimodular, and the same holds for matrix D. By a well known corollary of Hoffman and Kruskal's theorem (see p.268 in [12]), the optimum

$$max\{A_u\,\mathbf{q}_u \,|\, \mathbf{q}_u \geq 0 \quad \text{and} \quad D\,\mathbf{q}_u \leq \mathbf{e}\}$$

is reached (if it exists) at an integral point $\mathbf{x'}_u = \mathbf{x'}_f\,\mathbf{x'}_i$. Thus, under the conditions stated in the lemma, there exists an integer vector $\mathbf{x}' \geq 0$, namely $\mathbf{x}' = \mathbf{x}_c\,\mathbf{x'}_f\,\mathbf{x'}_i$, such that $A\,\mathbf{x}' \geq b + 1$, $-C\,\mathbf{x}' \leq M_0$, and $\mathbf{x}_f \leq \mathbf{x'}_f$. Clearly, $\mathbf{x}'$ is a forbidden vector, hence $\mathbf{x}$ is a suspect vector. ∎

Proposition 4.2 *The observable projections of the suspect vectors are the integral points of a polyhedron.*

Proof: Extending notations from lemma 4.1, let $\mathbf{r}_f = \mathbf{q}_f - \mathbf{x}_f$, $\mathbf{r}_i = \mathbf{q}_i$, and $\mathbf{r}_u = \mathbf{r}_f\,\mathbf{r}_i$. Then $\mathbf{x}_o = \mathbf{x}_c\,\mathbf{x}_f$ is the observable projection of a suspect vector iff there exists $\mathbf{r}_u \geq 0$ such that $-A_u\,\mathbf{r}_u \leq A_o\,\mathbf{x}_o - b - 1$ and $-C_u\,\mathbf{r}_u \leq M_0 + C_o\,\mathbf{x}_o$. By Farkas lemma, this system of inequalities has a non negative solution if and only if $y_a\,(A_o\,\mathbf{x}_o - b - 1) + \mathbf{y}_c\,(M_0 + C_o\,\mathbf{x}_o) \geq 0$ for all row-vectors $\mathbf{y} = y_a\mathbf{y}_c \geq 0$ in the cone $y_a\,A_u + \mathbf{y}_c\,C_u \leq 0$. The observable projections of the suspect vectors are therefore the integral points $\mathbf{x}_o$ of the polyhedron $(-y_a\,A_o - \mathbf{y}_c\,C_o)\,\mathbf{x}_o \leq \mathbf{y}_c\,M_0 - y_a(b+1)$ defined by the extremal rays of this cone. ∎

Denote by S the polyhedron constructed in Prop. 4.2, and let $H\,\mathbf{x}_o \leq \mathbf{k}$ be a linear definition of this *suspect* polyhedron, where matrix H and vector $\mathbf{k}$ are integer multiples of $(-y_a\,A_o - \mathbf{y}_c\,C_o)$ and $(\mathbf{y}_c\,M_0 - y_a(b+1))$, respectively.

Remarque 4.3 *The number of irredundant inequalities in $H\,\mathbf{x}_o \leq \mathbf{k}$, i.e. number of facets of S, is equal in the worst case to the number of extremal rays of the cone $y_a\,A_u + \mathbf{y}_c\,C_u \leq 0$, that grows exponentially with the size of matrix C_u. All the same, if one restricts to* integer *vectors $\mathbf{x}_o$, one can check whether $H\,\mathbf{x}_o \leq \mathbf{k}$ within time* polynomial *in the size of the marked graph G. Deciding whether an integer vector $\mathbf{x}_o = \mathbf{x}_c\,\mathbf{x}_f$ is the observable projection of a suspect vector amounts actually to deciding whether the linear sytem defined in lemma 4.1 has a solution in the* rational *vector variables $\mathbf{q}_f$ and $\mathbf{q}_i$.*

Now consider the transition system $K = (\mathcal{S}, T_o, \mathcal{T}, \mathbf{s}_0)$ defined as follows:

i) the initial state is the integer vector $\mathbf{s}_0 = -\mathbf{k}$,
ii) the set of states $\mathcal{S}$ is the subset of all integer vectors that may be reached from $\mathbf{s}_0$ by sequences of transitions in $\mathcal{T}$ (this set may be infinite),
iii) $T_o = T_c \cup T_f$ is the set of observable transitions of G,
iv) for any state $\mathbf{s} \in \mathcal{S}$ and for any uncontrollable transition $t \in T_f$, $\mathbf{s} \xrightarrow{t} \mathbf{s} + H\,\mathbf{t}$ is a transition in $\mathcal{T}$,
v) for any state $\mathbf{s} \in \mathcal{S}$ and for any controllable transition $t \in T_c$, $\mathbf{s} \xrightarrow{t} \mathbf{s} + H\,\mathbf{t}$ is a transition in $\mathcal{T}$ *iff* the target vector $\mathbf{s} + H\,\mathbf{t}$ has at least one positive entry.

If vector $\mathbf{k}$ has at least one negative entry, the most permissive policy enforcing constraint $A\,\mathbf{x} \leq b$ on plant G may be implemented by synchronizing (weakly) the plant and the controller K on the observable actions of the plant. Thus, when vector $\mathbf{x}$ has been fired in the plant, the controller is in a corresponding state $H\pi_o(\mathbf{x}) - \mathbf{k}$. In the converse case when vector $\mathbf{k}$ is non negative, the null vector $\mathbf{x} = 0$ is already suspect, and the control objective $A\,\mathbf{x} \leq b$ is just unfeasible.

The controller K defined above is nothing else than a finite vector addition system, whose firing rule consists in avoiding the non-negative state vectors. Notwithstanding, the dimension of the considered vectors may be very large, since it may be exponential in the size of G (see Remark 4.3). In order to reduce this dimension, one could, instead of deriving K from S, derive K from the integer hull S_1 of S (i.e., convex hull of all integral vectors in $\{\mathbf{q} \mid H\,\mathbf{q} \leq \mathbf{k}\}$), which is also a polyhedron. This would not suffice to break the exponential complexity of K, and would have little interest since computing S_1 from S has exponential cost.

The situation with respect to complexity is in fact not dramatic ar all. Relying on Lemma 4.1 and Remark 4.3 (second part), one may construct an equivalent controller $K = (\mathcal{S}, T_o, \mathcal{T}, \mathbf{s}_0)$ as follows:

i) $\mathcal{S}$ is the set of observable projections of firing vectors of G,
ii) $\mathbf{s}_0$ is the null vector,
iii) $T_o = T_c \cup T_f$ is the set of observable transitions of G,
iv) for any state $\mathbf{x}_o \in S$ and any uncontrollable transition $t \in T_f$, $\mathbf{x}_o \xrightarrow{t} \mathbf{x}_o + \mathbf{t}$ is a transition in $\mathcal{T}$,
v) for any state $\mathbf{x}_o \in S$ and for any controllable transition $t \in T_f$, $\mathbf{x}_o \xrightarrow{t} \mathbf{x}_o + \mathbf{t}$ is a transition in $\mathcal{T}$ iff the linear system with parameters $\mathbf{x}_c$ and $\mathbf{x}_f$ from lemma 4.1 has a rational solution in $\mathbf{q}_f$ and $\mathbf{q}_i$ at $\mathbf{x}_o = \mathbf{x}_c\,\mathbf{x}_f$.

The dimension of vectors is now linear in the size of G, and whether $\mathbf{x}_o \xrightarrow{t} \mathbf{x}_o + \mathbf{t}$ is a transition may be decided at each step in time polynomial in the size of G.

5 A Comparison with Monitors

This section is devoted to comparing our approach to the approach based on monitors, see e.g. [2], [14], or [7]. We shall establish a necessary and sufficient condition for the existence of monitor implementations of the most permissive control policy, and illustrate this condition with a counter-example. In contrast, our approach applies to all cases. It has also the advantage over the monitor approach to offer controllers with polynomial complexity. We will show that an exponential number of monitors is actually required, in the worst case, to implement the most permissive control policy when this is possible with monitors.

A pure monitor place p for plant G is defined with data as follows: an initial value $M_0[p] \in \mathbb{N}$, and a weight vector $C[p,\cdot\,] : T \to \mathbb{Z}$, subject to the restriction that $C[p,t\,] = 0$ for all unobservable transitions $t \in T_i$. For an observable transition $t \in T_o$, let $C[p,t] > 0$ mean an arc with the indicated weight from t to p, and $C[p,t] < 0$ mean an arc with the opposite weight from p to t. Adding to the plant a monitor place p has the neat effect of imposing on its firing vectors the linear constraint $M_0[p] + C[p,\cdot\,]\,\mathbf{x} \geq 0$. Adding several monitor places results in a conjunction of constraints. However, not all joint constraints are admissible: whenever $\mathbf{x}$ and $\mathbf{x} + \mathbf{t}$ are firing vectors and the latter does not satisfy the joint constraint, this should also apply to $\mathbf{x}$ if t is uncontrollable ($t \in T_u$).

Since monitor places are not connected to unobservable transitions, the constraints they impose bear only upon observable projections of firing vectors. Therefore, we focus from now on monitors given by an initial value $M_0[p]$ and a weight vector $C[p,\cdot\,] : T_o \to \mathbb{Z}$, and we consider exclusively the observable projections of the firing vectors.

From Prop. 2.10, the observable projections of the firing vectors of G are the integer points of a polyhedron F including the suspect polyhedron S. It may occur that some integer vectors in $F \setminus S$, hence not suspect, are not projections of vectors that may be fired in G under the most permissive control policy, because all paths reaching them are cut. Let V (for *viable*) denote the subset of integer vectors in $F \setminus S$ that can actually be fired under the most permissive control policy. Let W (for *wrong*) denote the set of the suspect vectors which may be reached from viable vectors in one step, i.e. vectors $(\mathbf{x} + \mathbf{t}) \in S$ such that $\mathbf{x} \in V$ and t is a (controllable) transition of G.

Theorem 5.1 *Assume $\mathbf{x} = 0$ is a viable vector. The most permissive control policy may be implemented by a set of* pure *monitor places (without self-loops) if and only if the convex hull of V does not intersect W. Moreover, it may be implemented by a single pure monitor place if and only if the convex hulls of V and W do not intersect.*

Proof:

The convex hull of V may be defined by a (finite or infinite) set of linear inequalities $\{\mathbf{a}_i\mathbf{x} \leq b_i \,|\, i \in I\}$, where vectors $\mathbf{a}_i$ are integer vectors and scalars b_i are integer numbers. As $\mathbf{x} = 0$ is a viable vector, all scalars b_i are non-negative.

Assume this convex hull does not intersect W. Let each inequality $\mathbf{a}_i\mathbf{x} \leq b_i$ be implemented by a monitor place p_i, with $M_0[p_i] = b_i$ and $C[p_i,\cdot\,] = -\mathbf{a}_i$. Thus, $M_0[p_i] + C[p_i,\cdot\,]\,\mathbf{x}$ is always non-negative for viable vectors $\mathbf{x} \in V$, whereas it must be negative for some index

i for any wrong vector $\mathbf{x} \in W$. From Prop. 3.5, $\{p_i \mid i \in I\}$ is an admissible set of monitors, implementing the most permissive control policy.

Assume on the contrary that some wrong vector $\mathbf{w} \in W$ belongs to the convex hull of V, i.e. $\mathbf{w}$ is a finite linear combination of vectors $\mathbf{v_1}, \ldots, \mathbf{v_m} \in V$ with positive rational coefficients. For any monitor place p, if $M_0[p] + C[p, \cdot]\,\mathbf{w} < 0$, then necessarily, $M_0[p] + C[p, \cdot]\,\mathbf{v}_j < 0$ for some viable vector $\mathbf{v}_j \in V$, and the most permissive control policy cannot be implemented by pure monitors.

Let us establish the second part of the theorem. Assume the most permissive control policy may be implemented by a single pure monitor place p. Then, $M_0[p] + C[p, \cdot]\,\mathbf{v} \geq 0$ for all viable vectors $v \in V$, while $M_0[p] + C[p, \cdot]\,\mathbf{w} \leq -1$ for all wrong vectors $\mathbf{w} \in W$. Therefore, V and W, as well as their convex hulls, are separated by the hyperplane $M_0[p] + C[p, \cdot]\,\mathbf{x} = -1/2$. Conversely, assume that convex hulls of V and W are disjoint, then they are separated by some hyperplane $\mathbf{ax} = b$, where $\mathbf{a}$ is an integer vector and b is a positive integer, and the most permissive control policy may be implemented by a single pure monitor place p, such that $M_0[p] = b$ and $C[p, \cdot] = -\mathbf{a}$. ∎

Remarque 5.2 *When the most permissive control policy can be implemented by a set of pure and bounded monitors, it can always be implemented by a finite set of pure and bounded monitors [1]. It is not clear to us whether a similar property of compactness holds for pure and unbounded monitors.*

In order to show that the first condition stated in Theo. 5.1 is not trivial, we propose a naive counter-example.

Example 5.3 *Consider the marked graph G with set of nodes $T = \{a, b, c\}$ and two unmarked arcs leading, respectively, from a and b to c. Clearly, G is a live marked graph. Assume that a and b are controllable and c is unobservable. Let the game be to prevent G from firing c, i.e. to enforce the constraint $\mathbf{x}[c] \leq 0$ on firing vectors of G. The set V of viable vectors is the set of vectors $\mathbf{y} \in \mathbb{N} \times \mathbb{N}$, counting the occurrences of a and b, such that $y[a] = 0$ or $y[b] = 0$. The convex hull of this set contains all vectors in $\mathbb{N} \times \mathbb{N}$, and in particular the wrong vector $\mathbf{y} = (1, 1)$. The most liberal control policy that prevents from firing c can therefore not be implemented with pure monitors.*

In order to show that the second condition stated in Theo. 5.1 is not equivalent to the first condition, we establish now a theorem that relies on a less naive counter-example. Recall that a marked graph is *strongly connected* if the underlying graph is strongly connected (i.e. if there exists a directed path from every node to every other node).

Theorem 5.4 *Let G be a safe and strongly connected marked graph. Suppose all transitions are observable, but some are uncontrollable. Given a linear constraint $AM \leq b$ on the markings of G, the most permissive control policy enforcing this constraint may always be implemented by an admissible set of pure monitor places. However, even in the restricted case of live and safe strongly connected marked graph, the number of monitor places required in the worst case grows exponentially with the number of transitions.*

Proof: The first statement in the theorem was proved in [2] for the larger class of safe and conservative Petri nets and for arbitrary forbidden state problems. In order to prove the second statement, let us consider the marked graph shown in Figure 1. The initial marking M_0 is such that all places with names p_i^j or q^j are unmarked and all places without a name are marked by 1. This marked graph is strongly connected, and it is both live and safe. All transitions represented as black nodes are controllable, all other transitions are uncontrollable (in particular t_{n+1}^1). The linear constraint to enforce on markings is $M[q_1] + M[q_2] \leq 1$.

Since this mark graph is safe and each place p_i^j or q^j has a complementary place, reachable markings may be identified with their induced restrictions on named places. Moreover, if one decomposes the marked graph into subnets N_i as shown in Fig. 1, a reachable marking may be identified with a family of $0,1$-vectors $M(N_i)$, with respective dimensions 3 for $M(N_0)$, 2 for $M(N_{n+1})$, and 6 for the other vectors $M(N_i)$. Let $\mathcal{M}$ be the subset of markings M represented in this way such that $M(N_0) = (0,0,1)$, $M(N_{n+1}) = (0,0)$, and, for $1 \leq i \leq n$, $M(N_i) = (x_1, x_2, x_3\,;\, x_4, x_5, x_6)$ is in the set

$$\{(0,1,1\,;\,1,0,1), (1,0,1\,;\,1,1,0), (1,1,0\,;\,0,1,1)\}$$

Each marking $M \in \mathcal{M}$ is reachable: $M_0[Y\rangle M$ where Y is the firing vector with projections $Y(N_i)$ as follows (disregarding transition t_{n+1}^1 that is not fired: $Y(N_0) = (1,1,1)$, $Y(N_{n+1}) = (0,0)$, and for remaining subnets, $Y(N_i) = \rho(M(N_i))$ where

$$\rho(x_1, x_2, x_3\,;\, x_4, x_5, x_6) = (x_1 + x_2 + x_3, x_2 + x_3, x_3\,;\, x_4 + x_5 + x_6, x_5 + x_6, x_6)$$

As t_{n+1}^1 cannot be fired twice without firing the controllable transition t_0^1, each marking $M \in \mathcal{M}$ is a viable marking, however it enables t_0^1 whose firing could lead to violate the constraint. Hence t_0^1 must be control disabled at every marking in $\mathcal{M}$. We show below that t_0^1 cannot be disabled by a single pure monitor place at two different markings in $\mathcal{M}$. As $card(\mathcal{M}) = 3^n$ and $card(T) = 6(n+1)$, the theorem will be proved.

Let M' and M'' be two markings in $\mathcal{M}$, reached respectively by firing vectors Y' and Y''. Assume $M'(N_i) \neq M''(N_i)$ for some i, and put

$$M'(N_i) + M''(N_i) = (y_1, y_2, y_3\,;\, y_4, y_5, y_6)$$

Thus, both vectors (y_1, y_2, y_3) and (y_4, y_5, y_6) belong to set $\{(2,1,1), (1,2,1), (1,1,2)\}$. Therefore, $M'(N_i) + M''(N_i)$ may be rewritten as the sum $\overline{M}'(N_i) + \overline{M}''(N_i)$ of respective markings

$$\overline{M}'(N_i) = (x_1', x_2', x_3'\,;\, x_4', x_5', x_6') \quad \text{and} \quad \overline{M}''(N_i) = (x_1'', x_2'', x_3''\,;\, x_4'', x_5'', x_6'')$$

such that $(x_1', x_2', x_3') = (x_4'', x_5'', x_6'') = (1,1,1)$ and both vectors (x_4', x_5', x_6') and (x_1'', x_2'', x_3'') are in the set $\{(1,0,0), (0,1,0), (0,0,1)\}$.

For $j \neq i$, let $\overline{M}'(N_j)$ and $\overline{M}''(N_j)$ be defined as above if $M'(N_j) \neq M''(N_j)$, or be equal

to $M'(N_j)$ and to $M''(N_j)$ if the latter are equal (in particular for $j = 0$ and $j = n+1$). It is crucial here to observe that even though transition t_0^1 happens to be fired from marking $\overline{M}'$ or from marking $\overline{M}''$, this does not enable to fire further any uncontrollable sequence of transitions that would lead to violate the constraint: in either case, one path from t_i^1 to t_{n+1}^1 or from t_i^4 to t_{n+1}^1 is marked with a single token, and both t_i^1 and t_i^4 are controllable.

Let firing vectors $Y' = \rho\ (M')$, $\overline{Y}' = \rho\ (\overline{M}')$, $Y'' = \rho\ (M'')$, and $\overline{Y}'' = \rho\ (\overline{M}'')$. One can easily verify that $M_0[Y'\rangle M'$, $M_0[\overline{Y}'\rangle \overline{M}'$, $M_0[Y''\rangle M''$, and $M_0[\overline{Y}''\rangle \overline{M}''$. Moreover, $\overline{Y}' + \overline{Y}'' = Y' + Y''$ follows from $\overline{M}' + \overline{M}'' = M' + M''$ by definition of ρ. These facts can be used to complete the proof.

Reasoning by contradiction, suppose actually that a single pure monitor place p is enough for disabling transition t_0^1 at both markings M' and M'' (without preventing to fire viable vectors Y' and Y''). Thus, if the considered monitor place has the initial value $M_0[p]$ and the weight vector $C[p, \cdot]$, it must be the case that $C[p, t_0^1] = -k$ with $k > 0$ such that

$$M_0[p] + C[p, \cdot]Y' < k \quad \text{and} \quad M_0[p] + C[p, \cdot]Y'' < k$$

On the other hand, transition t_0^1 should not be prevented to fire at markings $\overline{M}'$ and $\overline{M}''$: this would not be in agreement with the most liberal control policy. As a consequence,

$$M_0[p] + C[p, \cdot]\overline{Y'} \geq k \quad \text{and} \quad M_0[p] + C[p, \cdot]\overline{Y''} \geq k$$

A contradiction between the two assertions follows from $\overline{Y}' + \overline{Y}'' = Y' + Y''$ ∎

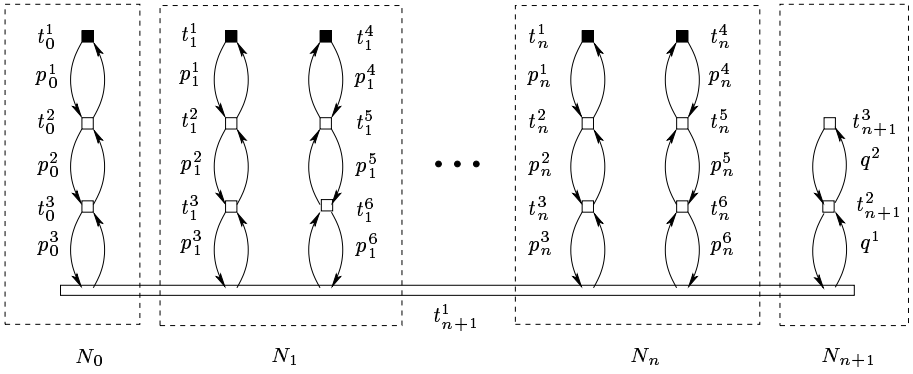

Figure 1: A safe, live, and strongly connected marked graph

6 Liveness preserving control of strongly connected marked graphs

A main limitation of linear controllers offered in section 4 is their incapacity to preserve liveness: installing control on plant G may result in the loss of liveness. This may be considered as a serious problem, all the more because the technical development relies heavily on the assumption of liveness of the original plant. The different techniques proposed in [13] and in [4] to enforce liveness in Petri nets cannot help to solve this problem, as the controlled plants we obtain can generally not be modelled as Petri nets. We do not know any solution in the general case. Therefore, we shall consider in this section the favourable case when plant G is given by a strongly connected live marked graph. The main contribution of sections 6 to 8 is to reduce in this case the liveness preserving control problem for marked graphs with uncontrollable / unobservable transitions to a classical problem on finite automata.

A strongly connected marked graph may be covered with circuits (there exists a directed path from every node to every other node), hence no transition can be a source node or an end node. As the number of tokens on each circuit is invariant under firing, the set of reachable markings is finite. As there exists a path from every node to every other node, the maximum firing deviation between two transitions (the maximum number of times node t or t' can fire without firing the other) is finite. Thus, a strongly connected marked graph is live if and only if it is deadlock-free.

We assume from now on that the plant taken into consideration is a live and strongly connected marked graph G. Therefore, the unrestricted behaviour of the plant is a rational language (because G is a bounded net), and every *non-blocking* controller preserves the liveness of the plant. Hence, if one could show that the set of firing sequences σ that are mapped to suspect vectors $\psi(\sigma)$ is a rational language, one could apply the iterative techniques proposed in [10] for computing non-blocking controllers. This indicates roughly the route taken in this section. Note that the rationality of the considered language does not follow directly from the polyhedral definition of the set of suspect vectors (since, e.g., the language $\{\sigma \in \{a,b\}^* \mid \psi(\sigma)(a) - \psi(\sigma)(b) \leq 0\}$ is *not* rational).

First, let us simplify the problem by getting rid of the unobservable actions. From Theo. 2.11 (corollary 2.12), the observable projections of the firing vectors (firing sequences) of G are the firing vectors (firing sequences) of a live marked graph G'. As the maximum firing deviation between two observable transitions in G is finite, and this deviation is preserved by projection, the same holds in G'. Therefore, G' is a strongly connected live marked graph. Now, the problem is to control the simplified plant G' so as to prevent this plant from firing any vector in the suspect polyhedron S (defined in Prop. 4.2), while keeping it live.

In order to alleviate the notation, let $G = (P, T, C, M_0)$ denote henceforth the simplified plant with set of transitions $T = T_c \cup T_f$. All transitions are observable, T_c is the subset of *controllable* transitions, and T_f is the subset of *free* transitions. Variables $\mathbf{x}$ and $\mathbf{q}$ range over $T \to \mathbb{N}$ and $T \to \mathbb{Q}$, respectively, and $\mathbf{1}$ is the constant vector with all entries $\mathbf{1}[t]$ equal

to 1. Finally, let $H\,\mathbf{q} \leq \mathbf{k}$ be a fixed linear system, where H is an integer matrix and $\mathbf{k}$ is an integer vector, defining the *suspect* polyhedron $S \subseteq (T \to \mathbb{Q}_+)$.

Given that no node in G is a source node or an end node, $C\,\mathbf{1} = 0$. Therefore, the rational polyhedron $F = \{\mathbf{q} \geq 0 \mid -C\,\mathbf{q} \leq M_0\}$, viz. the convex hull of the firing vectors of G, is closed under all translations that map vector $\mathbf{q}$ to $\mathbf{q} + \alpha\mathbf{1}$ for some scalar $\alpha \in \mathbb{Q}_+$. Moreover, no other translation maps F into F. Actually, if $C\,\mathbf{q} = 0$, then, for any two nodes t and t' connected by arc p in G, $\mathbf{q}[t] = \mathbf{q}[t']$ follows from $C[p,\cdot]\,\mathbf{q} = 0$, and this relation extends, by connectedness of G, to arbitrary pairs of nodes. Therefore, as $S \subseteq F$, the vector $\mathbf{v} = \mathbf{1}$ (or some positive multiple) is the unique vector likely to be a ray of the characteristic cone of S (recall that every polyhedron decomposes into the sum of a polytope, i.e. finite polyhedron, and a cone, called its characteristic cone). So, either $\mathbf{v} = \mathbf{1}$ is not a ray of the cone, and S is finite, or it is a ray of the cone, and S is a cylinder truncated in the lower part. It is easy to decide what case is met by applying standard linear programming algorithms. If S is finite, one may compute the maximal norm $\|q\| = \sum_j |q[j]| = \sum_j q[j] = q\,\mathbf{1}$ of the vectors $q \in S$. If S is infinite, one may compute the maximal norm of the vectors $\mathbf{q} \in S$ such that $\mathbf{q} - \mathbf{1} \notin S$. This is the maximum, for j ranging over rows of H, of the numbers

$$\mathbf{max}\,\{\mathbf{q}\,\mathbf{1} \mid \mathbf{q} \geq 0\,,\, H\,\mathbf{q} \leq \mathbf{k}\,,\, \mathbf{q}[j] < 1\}$$

$$\mathbf{max}\,\{\mathbf{q}\,\mathbf{1} \mid \mathbf{q} \geq \mathbf{1}\,,\, H\,\mathbf{q} \leq \mathbf{k}\,,\, H[i,\cdot]\,(\mathbf{q} - \mathbf{1}) > \mathbf{k}[i]\}$$

which may be either computed, or shown undefined, using standard algorithms. A similar norm may be computed for F. Let L be the smallest integer strictly larger than the two norms. The liveness preserving control problem may be now divided into two subproblems, one concerning firing vectors with norm less than L (*transient phase*), and the other concerning firing vectors with norm greater than L (*permanent phase*).

7 The Permanent Phase

We reduce in this part the liveness preserving control problem for the permanent phase to a similar problem about finite automata.

If S is finite, no control is needed for the permanent phase, hence we may assume that S is infinite. The idea is to project simultaneously the two cylinders $\{\mathbf{q} \in F \mid \mathbf{q}\,\mathbf{1} \geq L\}$ and $\{\mathbf{q} \in S \mid \mathbf{q}\,\mathbf{1} \geq L\}$ on an arbitrary hyperplane orthogonal to vector $\mathbf{1}$. One obtains in this way two polytopes F' and S' included in one another ($S' \subseteq F'$), such that all firing vectors $\mathbf{x} \in F$ project to a finite number of points $\phi(\mathbf{x}) \in F'$. The infinite trellis of all firing vectors $\mathbf{x} \in F$ projects thus to a finite automaton, such that transition $\phi(\mathbf{x}) \xrightarrow{t} \phi(\mathbf{x}')$ in the automaton represents relation $\mathbf{x}' = \mathbf{x} + \mathbf{t}$ in the trellis, and all transitions from $F' \setminus S'$ to S' result from controllable actions.

Before defining the folding morphism ϕ, let us simplify the expression of the set of suspect vectors $S_{\geq L} = \{\mathbf{q} \,|\, \mathbf{q}\,\mathbf{1} \geq L, H\,\mathbf{q} \leq \mathbf{k}\}$. Observe first that $H\,\mathbf{1} \leq 0$, since otherwise S would be bounded. Observe next that, for every non-negative vector $\mathbf{q}$ with norm $\|\mathbf{q}\| = \mathbf{q}\,\mathbf{1} \geq L$, it follows directly from the definition of L that $H\,\mathbf{q} \leq \mathbf{k}$ if and only if $H\,(\mathbf{q}+\mathbf{1}) \leq \mathbf{k}$. Therefore, whenever $H[i,\cdot]\,\mathbf{1} < 0$, the constraint $H[i,\cdot]\,\mathbf{q} \leq \mathbf{k}[i]$ is superfluous. By eliminating the superfluous constraints, one may obtain an equivalent expression of the set of suspect vectors $S_{\geq L} = \{\mathbf{q} \,|\, H''\,\mathbf{q} \leq \mathbf{k}'\}$ such that $H''\,\mathbf{1} = 0$. This expression is as follows.

Definition 7.1 *Let $H''\,\mathbf{q} \leq \mathbf{k}'$ be the linear system formed of all inequalities $H[i,\cdot]\,\mathbf{q} \leq \mathbf{k}[i]$ such that $H[i,\cdot]\,\mathbf{1} = 0$, plus inequalities $-\mathbf{q} \leq 0$ and $-\mathbf{q}\,\mathbf{1} \leq -L$.*

We can now come to the main definitions and propositions of the section. In the sequel, the columns of the connectivity matrix C are indexed by transitions in $T = \{t_1, \ldots, t_n\}$ in this order of enumeration. For any $\mathbf{v} \in \mathbb{Q}^{n-1}$, let $(\mathbf{v}, 0)$ denote vector $\mathbf{v}' \in \mathbb{Q}^n$ such that $\mathbf{v}'[j] = \mathbf{v}[j]$ for $j \in \{1, \ldots, n-1\}$ and $\mathbf{v}'[n] = 0$.

Definition 7.2 *Let $\phi : \mathbb{Q}^n \to \mathbb{Q}^{n-1}$ be the linear transformation $\phi(\mathbf{q}) = M\,\mathbf{q}$ defined by the $(n-1 \times n)$ matrix $M = (I, -\mathbf{1})$, in pictorial form:*

$$M = \begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 & 0 & -1 \\ 0 & 1 & 0 & \cdots & \cdots & 0 & 0 & -1 \\ \vdots & & \ddots & & & & \vdots & \vdots \\ \vdots & & & \ddots & & & \vdots & \vdots \\ \vdots & & & & \ddots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & 1 & 0 & -1 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 1 & -1 \end{pmatrix}$$

According to the above definition, $\phi(\mathbf{q})[j] = \mathbf{q}[j] - \mathbf{q}[n]$ for $j \in \{1, \ldots, n-1\}$.

Proposition 7.3 *$C\,\mathbf{q} = C\,(\phi(\mathbf{q}), 0)$ and $H''\,\mathbf{q} = H''\,(\phi(\mathbf{q}), 0)$.*

Proof: As $C\,\mathbf{1} = 0$, $C\,\mathbf{q} = C\,(\mathbf{q} - \mathbf{q}[n]\,\mathbf{1}) = C\,(\phi(\mathbf{q}), 0)$. As $H''\,\mathbf{1} = 0$, the same reasoning applies to the second part of the proposition. ∎

Definition 7.4 *Let $F' = \{\mathbf{q}' \,|\, -C'\,\mathbf{q}' \leq M_0\}$ and $S' = \{\mathbf{q}' \,|\, H'\,\mathbf{q}' \leq \mathbf{k}'\}$ where C' and H' are obtained by suppressing last column of C and H'', respectively.*

From Prop. 7.3, $F' = \phi(F)$ and $S' = \phi(S_{\geq L})$, hence $S \subseteq F$ entails $S' \subseteq F'$.

Proposition 7.5 *F' is a polytope.*

Proof: Suppose not, then $-C'\,\mathbf{q}' \leq 0$ for some vector $\mathbf{q}' \neq 0$. Given any rational number α, let $\mathbf{q}$ be the n-vector defined with $\mathbf{q}[n] = \alpha$ and $\mathbf{q}[j] = \mathbf{q}'[j] + \alpha$ for $j \in \{1, \ldots, n-1\}$. Thus $\mathbf{q}' = \phi(\mathbf{q})$. By Prop. 7.3, $-C\,\mathbf{q} \leq 0$, hence $\mathbf{q}$ is a ray of the characteristic cone of F. Therefore, $\mathbf{q}$ must be a multiple of vector $\mathbf{1}$, contradicting $\mathbf{q}' \neq 0$. ∎

Proposition 7.6 *The integer points of F' are the images under ϕ of the firing vectors of G.*

Proof: If $\mathbf{x}$ is a firing vector of G, then $\phi(\mathbf{x}) = M\,\mathbf{x}$ is clearly an integer vector. Conversely, let $\mathbf{z} \in \mathbb{Z}^{n-1}$ such that $-C'\,\mathbf{z} \leq M_0$, then $\mathbf{z} = \phi(\mathbf{z}, 0)$ and $-C'\,\mathbf{z} = -C\,(\mathbf{z}, 0)$. If $\mathbf{z} \geq 0$, $\mathbf{x} = (\mathbf{z}, 0)$ is a firing vector of G. Otherwise, let $\mathbf{x} = (\mathbf{z}, 0) + m\mathbf{1}$ where $-m$ is the least entry of vector $\mathbf{z}$, then $\mathbf{z} = \phi(\mathbf{x})$, and $-C\,\mathbf{x} = -C\,(\mathbf{z}, 0) = -C'\,\mathbf{z} \leq M_0$, so $\mathbf{x}$ is a firing vector. ∎

We can now extract from F' the automaton we were looking for.

Definition 7.7 (folding) *Let $\mathcal{A} = (\mathcal{S}, T, \mathcal{T}, \mathcal{S}_0)$ be the automaton as follows:*
– $\mathcal{S}$ is the set of all integer points in F',
– $T = \{t_1, \ldots, t_n\}$,
– for $t \in \{t_1, \ldots, t_{n-1}\}$, $\mathbf{z} \xrightarrow{\mathbf{t}} \mathbf{z}'$ is a transition in $\mathcal{T}$ iff $\mathbf{z}' = \mathbf{z} + \mathbf{t}$,
– for $t = t_n$, $\mathbf{z} \xrightarrow{\mathbf{t}} \mathbf{z}'$ is a transition in $\mathcal{T}$ iff $\mathbf{z}' = \mathbf{z} - \mathbf{1}$,
– $\mathbf{z}$ is initial ($\mathbf{z} \in \mathcal{S}_0$) iff $\mathbf{z} = \phi(\mathbf{x})$ for some firing vector with norm $\|\mathbf{x}\| = L$.

Proposition 7.8 *Let $\mathbf{x}$ be a firing vector ($\mathbf{x} \in F$) and let $\mathbf{z} = \phi(\mathbf{x})$. For any transition $t \in T$, $\mathbf{z} \xrightarrow{\mathbf{t}} \mathbf{z}'$ for some $\mathbf{z}' \in \mathcal{S}$ if and only if $\mathbf{x}' = \mathbf{x} + \mathbf{t}$ defines a firing vector ($\mathbf{x}' \in F$), and then $\mathbf{z}' = \phi(\mathbf{x}')$.*

Proof: Assume that $\mathbf{x}'$ is a firing vector. As $\mathbf{x}' = \mathbf{x} + \mathbf{t}$, it follows by linearity that $\phi(\mathbf{x}') = \phi(\mathbf{x}) + \phi(\mathbf{t})$. By definition 7.2, $\phi(\mathbf{t}) = \mathbf{t}$ if $t \in \{t_1, \ldots, t_{n-1}\}$ (where the former vector $\mathbf{t}$ has dimension n while the latter has dimension $n-1$), and $\phi(\mathbf{t}) = -\mathbf{1}$ if $t = t_n$. Thus, $\mathbf{z} \xrightarrow{\mathbf{t}} \mathbf{z}'$ in the automaton $\mathcal{A}$.
Assume now that $\mathbf{z} \xrightarrow{\mathbf{t}} \mathbf{z}'$ in the automaton $\mathcal{A}$. If $t \in \{t_1, \ldots, t_{n-1}\}$, then $\mathbf{z}' = \mathbf{z} + \mathbf{t} = \phi(\mathbf{x}) + \phi(\mathbf{t}) = \phi(\mathbf{x} + \mathbf{t})$, hence $-C\,\mathbf{x}' = -C'\,\mathbf{z}' \leq M_0$, and $\mathbf{x}'$ is a firing vector. If $t = t_n$, then $\mathbf{z}' = \mathbf{z} - \mathbf{1} = \phi(\mathbf{x}) + \phi(\mathbf{t}) = \phi(\mathbf{x} + \mathbf{t})$, hence $\mathbf{x}'$ is a firing vector for similar reasons. ∎

It follows from this proposition that every transition $\mathbf{z} \xrightarrow{\mathbf{t}} \mathbf{z}'$ of $\mathcal{A}$ with source $\mathbf{z} \in F' \setminus S'$ and target $\mathbf{z}' \in S'$ is labelled with a controllable action $t \in T_c$. Now, from Prop. 7.5, $\mathcal{A}$ is a finite automaton. Therefore, if we consider all non-blocking controllers, able to prevent $\mathcal{A}$ from reaching all suspect states $\mathbf{z} \in S'$ by disallowing controllable transitions or initial states or both, there exists among these one most liberal controller, as shown in [10]. This controller may be computed in the form of a finite automaton K_{perm}. It follows, by Prop.7.8, that K_{perm} is the most liberal controller able to enforce the control objective $H\,\mathbf{x} \not\leq \mathbf{k}$ on firing vectors of G in the permanent phase while preserving the liveness of this plant.

Since F' and S' are polytopes, it may look surprising that did not take advantage of geometry for computing the controller. This would probably not diminish complexity. In principle, one could avoid constructing automaton $\mathcal{A}$, and reason instead on finite unions of polytopes intersected with $\mathbb{Z}^{n-1}$. Namely, assuming that transition t_n is controllable, one might consider the following inductive definitions, where $\mathbf{z}$ ranges over $\mathbb{Z}^{n-1}$:

– $S'_0 = \{\mathbf{z} \,|\, \mathbf{z} \in S'\}$,
– $\Delta_i = \{\mathbf{z} \in F' \,|\, (\forall t) \;\; \mathbf{z} + \phi(\mathbf{t}) \in F' \Rightarrow \mathbf{z} + \phi(\mathbf{t}) \in S'_i\}$,
– $S'_{i+1} = S'_i \cup \{\mathbf{z} \in F' \,|\, (\exists \mathbf{z}' \in \Delta_i) \;\; \mathbf{z}_c = \mathbf{z'}_c \text{ and } \mathbf{z}_f \leq \mathbf{z'}_f\}$.
Computing iteratively S'_i would certainly converge in a finite number of steps to the set of states which the controller should bar access to. However, intersecting polytopes with $\mathbb{Z}^{n-1}$ is rather problematic for practical efficiency. Investigating another, non iterative, way of computing non-blocking controllers for automata with geometric representation, will be a subject of further research.

8 The transient phase

In order to complete the construction of the controller, we deal finally with the firing vectors $\mathbf{x}$ with norm $\|\mathbf{x}\| \leq L$. One may distinguish in this set suspect vectors such that $H\,\mathbf{x} \leq \mathbf{k}$. One may further distinguish in the subset of all vectors with norm L those vectors $\mathbf{x}$ such that $\phi(\mathbf{x})$ is not filtered out by controller K_{perm} from the initial states of $\mathcal{A}$. These *successful* vectors are not supect.

Now let automaton $\mathcal{B} = (Q, T, \mathcal{T}, Q_{init}, Q_{fin})$ with components as follows:
– Q is the set of firing vectors with norm $\|\mathbf{x}\| \leq L$,
– T is the set of transitions of G,
– $\mathbf{x} \xrightarrow{\mathbf{t}} \mathbf{x}'$ is a transition in $\mathcal{T}$ iff $\mathbf{x}' = \mathbf{x} + \mathbf{t}$,
– $Q_{init} = \{\mathbf{0}\}$,
– Q_{fin} is the subset of the successful vectors.
This automaton is finite and acyclic, hence the most permissive controller, able to ensure successful termination of $\mathcal{B}$ by disallowing controllable transitions when necessary, may be computed as a finite acyclic automaton K_{trans}. It remains to identify the final states of K_{trans} with the initial states of K_{perm} to obtain the expected controller for G, namely $K = K_{trans}\,;K_{perm}$. By synchronizing (weakly) plant G and controller K on the observable actions of the plant, one imposes actually on live marked graph G the most permissive control policy that enforces the linear constraint specified on the firing vectors of G while keeping all transitions of the plant live.

References

[1] Droste, M., Shortt, R.M.: Bounded Petri Nets of Finite Dimension Have Only Finitely Many Reachable Markings. Bulletin of the EATCS **48** (1992) 172–175

[2] Giua, A., Di Cesare, F., Silva, M.: Generalized Mutual Exclusion Constraints on Nets with Uncontrollable Transitions. Proc. IEEE-SMC (1992) 974–979

[3] Holloway, L., Krogh, B., Giua, A.: A Survey of Petri Net Methods for Controlled Discrete Event Systems. Discrete Event Dynamic Systems: Theory and Applications **7** (1997) 151–190

[4] Iordache, M., Antsaklis, P.: Generalized Conditions for Liveness Enforcement and Deadlock Prevention in Petri Nets. Proc. ICATPN, Springer-Verlag LNCS **2075** (2001) 184–203

[5] Li, Y., Wonham, W.M.: Control of Vector Discrete-Event Systems I – The Base Model. IEEE Trans. on Automatic Control **38** no.8 (1993) 1214–1227

[6] Li, Y., Wonham, W.M.: Control of Vector Discrete-Event Systems II – Controller Synthesis. IEEE Trans. on Automatic Control **39** no.3 (1994) 512–531

[7] Moody, J., Antsaklis, P.: Petri Net Supervisors for DES with Uncontrollable and Unobservable Transitions. IEEE Transactions on Automatic Control **45** no.3 (2000) 462–476

[8] Murata, T.: Circuit Theoretic Analysis and Synthesis of Marked Graphs. IEEE Transactions on Circuits and Systems **CAS-24** no.7 (1977) 400–405

[9] Murata, T.: Petri Nets: Properties, Analysis and Applications. Proceedings of the IEEE **77** no.4 (1989) 541–580

[10] Ramadge, W., Wonham, W.: On the Supremal Controllable Language of a Given Language. SIAM Journal of Control and Optimization **8** no.3 (1987) 637–659

[11] Rezg, N., Xie, X., Ghaffari, A.: Supervisory Control in Discrete Event Systems using the Theory of Regions. Discrete Event Systems: Analysis and Control, Kluwer (2000) 391–398

[12] Schrijver, A.: Theory of Linear and Integer Programming. John Wiley and Sons (1986)

[13] Valk, R., Jantzen, M.: The Residue of Vector Sets with Applications to Decidability Problems in Petri Nets. Acta Informatica **21** (1985) 643–674

[14] Yamalidou, K., Moody, J., Lemmon, M., Antsaklis, P.: Feedback Control of Petri Nets based on Place Invariants. Automatica **32** no.1 (1996) 15–28

Unité de recherche INRIA Rennes
IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399