Discrete Event Dynamic Systems (2025) 35:87–105
https://doi.org/10.1007/s10626-025-00412-x

Optimization-based computation of bounded sequences to reach target states in DESs

Roberto Cordone[1] · **Francesco Basile**[2] · **Luigi Piroddi**[3]

Received: 12 November 2024 / Accepted: 14 April 2025 / Published online: 23 April 2025

Abstract
The enumeration of legal transition paths leading to a target state (or set of states) is of paramount importance in the control of discrete event systems, but is hindered by the state explosion problem. A method is proposed in this paper, in the context of Petri nets, to calculate and enumerate firing count vectors for which there exists at least an admissible transition sequence leading to a given target marking. The method is shown to improve the approach based on singular complementary transition invariants proposed by Kostin and combines an integer linear programming formulation that finds the shortest minimal solution and a branching procedure that realizes a partition of the solution set. The enumeration can be restricted to minimal solutions or extended to non-minimal ones. Moreover, the approach is extended by adding a further constraint that the target transition sequences should pass by intermediate markings (in a specific order or not). Finally, source, target and via markings can be replaced by sets of markings. Some analytical examples are discussed in detail to show the effectiveness of the proposed approach.

Keywords Reachability analysis · Petri nets · Target marking · Waypoints · Mathematical programming

1 Introduction

An often encountered problem in the control of manufacturing systems –and discrete event systems, in general– is that of enumerating the feasible transition paths leading from the current state to a set of target states (possibly passing through intermediate states as well, as

---

✉ Francesco Basile
fbasile@unisa.it

Roberto Cordone
roberto.cordone@unimi.it

Luigi Piroddi
luigi.piroddi@polimi.it

[1] Dipartimento di Informatica, Università degli Studi di Milano, Via Celoria, 18, Milano, (MI) 20133, Italy

[2] Dipartimento di Ingegneria dell'informazione ed elettrica e matematica applicata, Università di Salerno, Fisciano, (SA) 84084, Italy

[3] Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Via Ponzio 34/5, Milano, (MI) 20133, Italy

often requested in navigation or logistic problems). Then, one such path is selected according to the criterion of interest (which could be related *e.g.* to cost, time, or efficiency) and enforced by an agent acting as controller, scheduler or supervisor. Since in various applications of interest such decisions have to be taken in real time (Basile et al. 2022), it is crucial to employ efficient algorithms to solve such reachability problems.

We investigate this issue for untimed Petri net (PN) models. The decidability of the reachability problem (*i.e.*, determining if a certain marking $\boldsymbol{m}_t$ is reachable from the initial marking $\boldsymbol{m}_0$) has been proven for generalized PNs in Mayr (1984), but the method suggested there essentially amounts to performing a full reachability analysis of the PN, an operation of exponential complexity which can require a prohibitive large computational effort in large scale problems. In general, all methods based on reachability analysis tend to suffer from the state explosion problem. This issue can be somewhat contained if a bound is posed on the exploration depth of the reachability tree, as done *e.g.* in Reveliotis (2005). Bounding the length of transition sequences is also justified in view of the fact that in a bounded system a state is reachable in a finite number of steps. In this respect, some results are available to estimate this number (Basile et al. 2012; Chouchane et al. 2023), that can be used to appropriately dimension the mentioned bound. Another possibility to contain the state explosion problem is to employ an abstraction, such as *e.g.* stubborn sets (Schmidt 1999), as a tool to reduce the state space while preserving certain system properties. Notice, however, that any such technique will also result in a reduction of the information conveyed by the reachability graph in terms of markings (see for example the coverability graph) or sequences (as stubborn sets do) while preserving a set of properties. For the purpose of this work, which focuses on deciding if a marking is reachable and how it can be reached, a complete representation of feasible sequences is required.

Alternative algebraic methods are based on the state equation, which relates the initial and target marking by way of the incidence matrix and the firing count vector (FCV). In this context, the reachability problem is reformulated as that of finding all FCVs resulting in a net marking variation equal to $\boldsymbol{m}_t - \boldsymbol{m}_0$. Enumerating FCVs as opposed to transition sequences is quite appealing as it typically entails a significant reduction in the number of solutions, given that the same FCV may encompass many transition sequences. However, the existence of a FCV satisfying the state equation does not immediately guarantee that there exist any *legal* firing sequences solving the reachability problem, since the state equation only accounts for the marking variations resulting from the firing of transitions, without ensuring their fireability. Another source of complexity stems from the fact that the number of solutions could be unbounded (an acceptable sequence can sometimes be infinitely extended, by interlacing it an arbitrary number of times with additional cyclic transition sequences resulting in a null marking variation). For the interested reader, a comprehensive review regarding reachability analysis with the state equation is provided in Su et al. (2023) and Qi et al. (2023).

In the works of Kostin (see, *e.g.*, Kostin 2003, 2006, 2008) the concept of *complemented* Petri net is introduced, where a transition is added to the net suitably connected to the net places so that its firing produces a net marking variation equal to $\boldsymbol{m}_0 - \boldsymbol{m}_t$. The T-invariants of the complemented net which include a single firing of the auxiliary transition (denoted *singular complementary* T-invariants, or SCTs) correspond to FCVs of the original net solving the reachability problem from $\boldsymbol{m}_0$ to $\boldsymbol{m}_t$. This comes in particularly handy, since various efficient techniques for the calculation of T-invariants can be found in the literature (see, *e.g.*, Martínez and Silva 1982; Colom and Silva 1991; Silva et al. 1998). Notice that all T-invariants of a Petri net can be generated by linear combination of the *minimal-support* T-invariants (Murata 1989), which in turn are a *finite* number (Kostin 2003). The reachability problem can be greatly simplified by calculating the SCTs and then restricting attention in the reachability analysis to the legal transition sequences that are compatible with them. Unfortunately, one cannot limit the analysis to a subset of the SCTs (*e.g.* the *minimal-support* SCTs), but has necessarily to take into account *all* SCTs, which are generally infinite. This follows directly upon noticing that the existence of SCTs is only a necessary condition for the existence of legal firing sequences solving the reachability problem. Therefore, it may well happen that no legal firing sequence is compatible with a minimal-support SCT, unless the latter is linearly combined with other T-invariants.

Various mechanisms leading to minimal and non-minimal SCTs, such as that of "token borrowing" among T-invariants, are discussed in Kostin (2006, 2008), but unfortunately they do not translate into an efficient and systematic search/enumeration algorithm. In this paper, we employ the general concept of SCTs (Kostin 2003) to derive a method for the enumeration of FCVs resulting in at least one admissible transition sequence (denoted admissible FCVs in the sequel, with a slight abuse of notation) starting from a source marking $\boldsymbol{m}_0$ and ultimately leading to a target marking $\boldsymbol{m}_t$.

However, if the objective is extended to include waypoints and non-singleton target sets, even the useful concept of SCT turns out to be impractical. Indeed, SCTs are by definition "source and target state dependent". In other words, they are defined with reference to a pair $(\mathbf{m}_0, \mathbf{m}_t)$. If for example more than one target marking $\mathbf{m}_{t_k}$ is considered, then the reachability analysis must be repeated for each pair $(\mathbf{m}_0, \mathbf{m}_{t_k})$, computing the relative SCTs (a similar reasoning applies to multiple source markings). Furthermore, if a waypoint $\boldsymbol{m}_w$ is introduced, all combinations of SCTs associated to paths from $\boldsymbol{m}_0$ to $\boldsymbol{m}_w$ and from $\boldsymbol{m}_w$ to $\boldsymbol{m}_t$ must be considered. The complexity becomes even worse if multiple successive waypoints are required, or if those too are expressed as marking sets rather than single markings, possibly leading to combinatorial explosion. This is a strong limitation of this and other reachability approaches based on forward or backward analysis. For example, in Su et al. (2023) and Qi et al. (2023) a backward algorithm is presented to determine the existence of a legal firing sequence for an FCV satisfying the state equation. While very efficient, this approach is "source and target state dependent" in the sense explained above, and therefore not suitable for the problem considered in this paper.

This paper exploits the state equation to build a suitable formulation of the reachability problem such that the solution is not limited to a single source and target state. This comes in the form of a single optimization problem that can be used to enumerate FCVs satisfying a number of complex reachability requirements, including waypoints and marking sets. The solutions are guaranteed to admit at least one admissible transition sequence (in the sequel we will denote such solutions as *admissible* FCVs). The presented approach is not restricted to a specific net subclass and is optimization-based. This last feature is particularly convenient if the reachability analysis is aimed at a specific application, such as control, scheduling or supervision, so that one path must be selected among the admissible ones, according to a criterion of interest (which could be related *e.g.* to cost, time, or efficiency). By formulating the overall problem in an optimization form, one can obtain the solution in a single stage.

The novel contribution can be summarized as follows:

- A systematic procedure is presented to enumerate all the admissible FCVs. The procedure operates on FCVs but at the same time performs an admissibility check.
- A practical approach is proposed to handle the problem that the number of solutions could be infinite (*e.g.*, by cycling indefinitely along T-invariants of the net), that consists in posing a bound K on the maximum number of allowed transition firings, as in Reveliotis (2005);

- The enumeration problem can be solved efficiently by combining a basic integer linear programming (ILP) formulation, that characterizes the FCVs satisfying the reachability requirements, with a Branch & Bound (B&B) procedure that efficiently partitions the solution space. This is well in line with a recent trend in PN-related research, which indicates that many analysis, control, and fault diagnosis problems can be more conveniently formulated and solved as optimization problems, usually in the form of ILP problems (see, *e.g.*, Basile et al. 2012; Cordone and Piroddi 2013; Cordone et al. 2013; Basile et al. 2015; Cong et al. 2018). Indeed, it turns out that, despite their computational complexity, *optimization-based* approaches can be practically more convenient when compared to alternative solutions, since they rely on *off-the-shelf* optimized solvers, as opposed to *ad hoc* algorithms. In particular, various efficient software suites can be employed to tackle ILP problems, such as CPLEX® or FICO™ Xpress.
- SCTs are employed to derive the basic formulation, but are not carried over to the more complex cases with multiple target markings or with waypoints, for computational convenience. Instead, the optimization model relies on the state equation, admissibility conditions and a suitable formulation of the constraints.

The rest of this manuscript is organized as follows. In Section 2 a reminder on the basic concepts and notation regarding Petri nets is provided, together with the essential notions related to the algebraic approach of Kostin to the reachability problem, based on the SCTs. Section 3 introduces the ILP formulation for the characterization of the admissible FCVs (up to a given length K) for various types of reachability specifications. The B&B scheme is illustrated next, in Section 4. Finally, various analytical examples are discussed in some detail in Section 5, before the concluding remarks.

2 Preliminaries

2.1 Petri nets

We assume that the reader is accustomed to the basic definitions and notation regarding PNs and only some essential properties will be recalled for ease of reference (see Murata 1989 for further details).

A marked PN is a 5-tuple $N = \langle P, T, \boldsymbol{Pre}, \boldsymbol{Post}, \boldsymbol{m}_0 \rangle$, where P and T are the (finite and nonempty) sets of n places and m transitions, with $P \cap T = \emptyset$, $\boldsymbol{Pre}, \boldsymbol{Post} \in \mathbb{N}^{n \times m}$ are the input and output matrices ($\mathbb{N}$ being the set of nonnegative integers), and $\boldsymbol{m}_0 \in \mathbb{N}^n$ is the (initial) marking vector. Places (graphically represented as circles) are connected to transitions (represented as bars) through directed weighted arcs. Places and arcs are endowed with integer values called tokens and weights, respectively. The marking vector $\boldsymbol{m}$ defines the distribution of tokens in places. The coefficient $Pre(k, j)$ $[Post(k, j)]$ represents the weight of an arc going from p_k $[t_j]$ to t_j $[p_k]$ (the coefficient is set to 0 if the arc is absent). The incidence matrix $\boldsymbol{C} = \boldsymbol{Post} - \boldsymbol{Pre}$ is often provided as an alternative to the pair $(\boldsymbol{Pre}, \boldsymbol{Post})$ to describe the PN topology (the two representations are equivalent in the absence of self-loops). A transition $t_j \in T$ is enabled in a marking $\boldsymbol{m}$ (denoted $\boldsymbol{m}[t_j\rangle$) iff $\boldsymbol{m} \geq \boldsymbol{Pre}\, \boldsymbol{\epsilon}^{(j)}$, where $\boldsymbol{\epsilon}^{(j)}$ is the jth versor of the $\mathbb{R}^m$ coordinate space (*i.e.*, $\epsilon_k^{(j)} = 1$ if $k = j$ and 0 otherwise). A transition t_j such that $\boldsymbol{m}[t_j\rangle$ may fire at marking $\boldsymbol{m}$, yielding the marking $\boldsymbol{m}'$ (denoted $\boldsymbol{m}[t_j\rangle \boldsymbol{m}'$), where $\boldsymbol{m}' = \boldsymbol{m} + \boldsymbol{C}\boldsymbol{\epsilon}^{(j)}$.

More in general, the effect of a sequence of transition firings can be computed in a single operation using the FCV $\boldsymbol{\sigma}$, whose element σ_j represents the number of times that transition

t_j fires in the sequence, using the state equation $\boldsymbol{m}' = \boldsymbol{m} + \boldsymbol{C}\boldsymbol{\sigma}$. A vector $\boldsymbol{\tau} \geq 0$ such that $\boldsymbol{C}\boldsymbol{\tau} = \boldsymbol{0}$ is called a T-invariant. An admissible transition sequence whose FCV coincides with a T-invariant produces a null net marking variation ($\boldsymbol{m}' = \boldsymbol{m}$), thus taking the PN back to the initial marking. The set of transitions $||\boldsymbol{\tau}|| = \{t_j \in T \mid \tau_j > 0\}$ is called the support of the T-invariant. A T-invariant $\boldsymbol{\tau}$ has minimal support if there does not exist another T-invariant $\boldsymbol{\tau}'$ such that $||\boldsymbol{\tau}'|| \subset ||\boldsymbol{\tau}||$. A T-invariant $\boldsymbol{\tau}$ is minimal if there does not exist another T-invariant $\boldsymbol{\tau}'$ such that $\boldsymbol{\tau}' \leq \boldsymbol{\tau}$. A *minimal-support* T-invariant has minimal support and is minimal. The set of such minimal-support T-invariants is finite and constitutes a basis: any T-invariant can be computed by linear combination of minimal-support T-invariants (Memmi and Roucairol 1980).

The reachability set $R(N, \boldsymbol{m}_0)$ collects the markings reachable from $\boldsymbol{m}_0$ by way of admissible transition sequences. The reachability graph is a directed graph $RG = (V, A)$, where $V = R(N, \boldsymbol{m}_0)$ is the set of vertices and $A \subseteq (V \times V)$ the set of arcs, associated with the PN transitions through a labeling function $h : A \rightarrow T$.

A place $p_i \in P$ is bounded iff $\exists k > 0$ s.t. $m_i \leq k, \forall \boldsymbol{m} \in R(N, \boldsymbol{m}_0)$. A PN is bounded iff all its places are bounded. A transition $t_j \in T$ is live iff $\forall \boldsymbol{m} \in R(N, \boldsymbol{m}_0)$, $\exists \boldsymbol{m}' \in R(N, \boldsymbol{m})$ s.t. $\boldsymbol{m}'[t_j\rangle$. N is live iff all its transitions are live. N is reversible iff $\forall \boldsymbol{m} \in R(N, \boldsymbol{m}_0)$, $\boldsymbol{m}_0 \in R(N, \boldsymbol{m})$. A marking $\boldsymbol{m} \in R(N, \boldsymbol{m}_0)$, s.t. $\nexists t_j \in T$ enabled in $\boldsymbol{m}$, is called a dead marking and represents a (total) deadlock state. A PN with no reachable dead markings is called deadlock-free.

2.2 An algebraic perspective of the reachability problem

Given a marked PN $N = \langle P, T, \boldsymbol{Pre}, \boldsymbol{Post}, \boldsymbol{m}_0\rangle$ and a target marking $\boldsymbol{m}_t$, one can define a *complemented* net (Kostin 2003) $N_c = \langle P, T \cup \{t_{m+1}\}, [\boldsymbol{Pre}\ \boldsymbol{m}_t], [\boldsymbol{Post}\ \boldsymbol{m}_0], \boldsymbol{m}_0\rangle$, by adding an auxiliary transition, t_{m+1}, suitably connected to the places of the original PN. The incidence matrix of N_c turns out to be $\boldsymbol{C}_c = [\boldsymbol{C}\ (\boldsymbol{m}_0 - \boldsymbol{m}_t)]$. We denote as $\mathcal{T} = \{\boldsymbol{\tau} \mid \boldsymbol{C}_c\boldsymbol{\tau} = \boldsymbol{0}\}$ the set of T-invariants and as $\mathcal{T}^{\text{ms}} \subseteq \mathcal{T}$ the set of minimal-support T-invariants of N_c.

The complemented net has the property that if t_{m+1} is fired in $\boldsymbol{m}_t$ it leads back to the initial marking $\boldsymbol{m}_0$. Indeed, $\boldsymbol{m}_t + \boldsymbol{C}_c\boldsymbol{\epsilon}^{(m+1)} = \boldsymbol{m}_t + (\boldsymbol{m}_0 - \boldsymbol{m}_t) = \boldsymbol{m}_0$. This implies that a necessary condition for the reachability of $\boldsymbol{m}_t$ from $\boldsymbol{m}_0$ in N is the existence of a T-invariant of N_c with a single firing of t_{m+1}.

T-invariants of N_c with this property are denoted SCTs (Kostin 2003). Instead, the T-invariants of N_c with no firing of t_c (denoted non-complementary T-invariants, or NCTs) correspond to the T-invariants of N (Kostin 2003). The remaining T-invariants of N_c are referred to as non-singular complementary T-invariants (NSCTs). Accordingly, we can partition $\mathcal{T}$ as $\mathcal{T} = \mathcal{T}_0 \cup \mathcal{T}_1 \cup \mathcal{T}_{2+}$, where $\mathcal{T}_0 = \{\boldsymbol{\tau} \in \mathcal{T} \mid \tau_{m+1} = 0\}$, $\mathcal{T}_1 = \{\boldsymbol{\tau} \in \mathcal{T} \mid \tau_{m+1} = 1\}$, and $\mathcal{T}_{2+} = \{\boldsymbol{\tau} \in \mathcal{T} \mid \tau_{m+1} \geq 2\}$ are three sets including the NCTs, the SCTs, and the NSCTs, respectively. We will further denote as $\mathcal{T}_1^a \subseteq \mathcal{T}_1$ the set of SCTs corresponding to admissible FCVs solving the reachability problem.

The utility of the SCT concept resides in the fact that the problem of finding fireable transition sequences leading to $\boldsymbol{m}_t$ can be addressed without a full reachability analysis, but focusing only on sequences compatible with SCTs. The first task to accomplish is therefore that of calculating such T-invariants. While there exist many efficient algorithms for the calculation of *minimal-support* T-invariants (which are in a finite number and can be used to generate all other T-invariants by linear combination), it is however arguable that one should limit the analysis to these T-invariants alone. Indeed, it sometimes happens that no fireable transition sequence corresponds to a minimal-support T-invariant, while legal transition sequences may be found if the T-invariant is linearly combined with other minimal-support T-invariants. As a consequence, we cannot rule out T-invariants with non-minimal support.

In Kostin (2003) it is argued that there exists only a finite number of SCTs that can be obtained by linear combination (with rational coefficients) of *minimal-support* SCTs and NSCTs (that is, not including the NCTs in the linear combination). This implies that they can be fully computed with some enumeration technique. However, this does not fully solve the problem, since each such SCT can be further linearly combined (with integer coefficients) with the NCTs, obtaining infinite SCTs.

3 Problem formulation

3.1 Problem 1: Reaching a target marking

The basic reachability problem consists in determining the admissible transition sequences starting from $\boldsymbol{m}_0$ and arriving to a target marking $\boldsymbol{m}_t$. As discussed before, we focus instead on enumerating the admissible FCVs associated with the mentioned legal transition sequences, to reduce the solution space to a more manageable size. Given the FCVs, the actual admissible transition sequences can be determined by standard reachability analysis, limiting the exploration to trajectories compatible with the FCVs.

For the purpose of enumerating said FCVs, we construct the complemented net N_c and use a standard method to compute its minimal-support T-invariants, *i.e.* the elements of $\mathcal{T}^{\text{ms}} = \{\boldsymbol{\tau}^{(1)}, \ldots, \boldsymbol{\tau}^{(p)}\}$. Now, according to Kostin (2003) any element $\boldsymbol{\tau} \in \mathcal{T}_1$ (recall that $\mathcal{T}_1$ is the set of SCTs) can be constructed as follows:

$$\boldsymbol{\tau} = \sum_{j=1}^{p} \alpha_j \boldsymbol{\tau}^{(j)} \tag{1}$$

where the coefficients $\alpha_j \geq 0$, $j = 1, \ldots, p$, are such that: i) all elements of $\boldsymbol{\tau}$ are integers, and ii) $\tau_{m+1} = \sum_{j=1}^{p} \alpha_j \tau_{m+1}^{(j)} = 1$. It follows from the integrality requirement that all the α_j coefficients must be rational numbers. Furthermore, the coefficients α_j associated with SCTs and NSCTs of $\mathcal{T}^{\text{ms}}$ must not exceed 1 for condition (ii) to hold. On the other hand, the coefficients α_j associated with NCTs of $\mathcal{T}^{\text{ms}}$ are not involved in condition (ii), and can thus be allowed to exceed 1.

Now, consider a generic transition sequence $(t_{j_1}, \ldots, t_{j_K})$ of length K, where $j_1, \ldots, j_K \in \{1, \ldots, m\}$. The corresponding firing vector can be written as $\boldsymbol{\sigma} = \sum_{k=1}^{K} \boldsymbol{\epsilon}^{(j_k)}$ where $\boldsymbol{\epsilon}^{(j)}$ is the jth versor of the $\mathbb{R}^m$ coordinate space. Thus, another representation of $\boldsymbol{\tau}$ can be obtained as follows:

$$\boldsymbol{\tau} = \sum_{j=1}^{p} \alpha_j \boldsymbol{\tau}^{(j)} = \begin{bmatrix} \boldsymbol{\sigma} \\ 1 \end{bmatrix}, \tag{2}$$

provided that $\boldsymbol{\sigma}$ solves the expression $\boldsymbol{m}_t = \boldsymbol{m}_0 + \boldsymbol{C}\boldsymbol{\sigma}$. This formulation, with $\boldsymbol{\sigma}$ constructed as a sum of versors, ensures the integrality of vector $\boldsymbol{\tau}$ (albeit limiting the length of the considered sequences to a number K). One important consequence of this fact is that it is not necessary to specify explicitly that parameters α_j should be rational, as this comes free of charge.

Another advantage of the previous formulation is that it can be used to enforce the fireability of the corresponding sequence by setting the following constraints:

$$\boldsymbol{m}_0 + \boldsymbol{C} \sum_{k=1}^{i-1} \boldsymbol{\epsilon}^{(j_k)} \geq \boldsymbol{Pre}\ \boldsymbol{\epsilon}^{(j_i)}, \quad i = 1, \ldots, K. \tag{3}$$

Notice that the set of markings $\boldsymbol{m}_t$ that can be obtained by way of the state equation with all possible FCVs of the form $\boldsymbol{\sigma} = \sum_{k=1}^{K} \boldsymbol{\epsilon}^{(j_k)}$ for which conditions Eq. 3 hold coincides with the set of reachable markings in K steps (Reveliotis 2005).

Assembling the previously developed expressions, we can obtain an ILP formulation characterizing the legal firing sequences solving the reachability problem from $\boldsymbol{m}_0$ to $\boldsymbol{m}_t$ (provided there is at least one such sequence that requires not more than K steps), together with the corresponding firing vectors (or, which is the same, the corresponding SCTs):

$$\begin{aligned}
&\text{minimize} \sum_{k=1}^{K} \sum_{j=1}^{m} e_j^{(k)} \\
&\text{subject to} \\
&\boldsymbol{m}_t = \boldsymbol{m}_0 + \boldsymbol{C}\boldsymbol{\sigma} \\
&\boldsymbol{\sigma} = \sum_{k=1}^{K} \boldsymbol{e}^{(k)} \\
&\begin{bmatrix} \boldsymbol{\sigma} \\ 1 \end{bmatrix} = \sum_{j=1}^{p} \alpha_j \boldsymbol{\tau}^{(j)} \\
&\boldsymbol{m}_0 + \boldsymbol{C} \sum_{k=1}^{i-1} \boldsymbol{e}^{(k)} \geq \boldsymbol{Pre}\ \boldsymbol{e}^{(i)} && i = 1, \ldots, K \\
&\sum_{j=1}^{m} e_j^{(k)} \leq 1 && k = 1, \ldots, K \\
&\sum_{j=1}^{m} e_j^{(k)} \geq \sum_{j=1}^{m} e_j^{(k+1)} && k = 1, \ldots, K-1 \\
&e_j^{(k)} \in \{0, 1\} && j = 1, \ldots, m,\ k = 1, \ldots, K \\
&\alpha_j \geq 0 && j = 1, \ldots, p
\end{aligned}$$

In the previous formulation, the column vector $\boldsymbol{e}^{(k)}$ (corresponding to the kth firing in the sequence) can either be a versor (thus accounting for an actual transition firing) or a null vector, given that its parameters are binary but their sum cannot exceed 1 (it can be either 0 or 1). To avoid a multiplicity of solutions corresponding to the same sequence with less than K transition firings, void transitions are forced to be at the end of the sequence (by imposing that the sum of the elements of $\boldsymbol{e}^{(k+1)}$ cannot exceed that of $\boldsymbol{e}^{(k)}$). The cost function ensures that the solution will be a fireable sequence with the shortest length among all the admissible solutions. In the sequel, we will denote the previous ILP formulation as a pair $(f_0, \mathcal{C}_0)$, f_0 denoting the cost function and $\mathcal{C}_0$ the set of constraints.

Notice that the set of admissible solutions $\boldsymbol{\tau} = [\boldsymbol{\sigma}^T\ 1]^T$ for the set of constraints $\mathcal{C}_0$ equals $\mathcal{T}_1^{\star} = \mathcal{T}_1^a \cap \{\boldsymbol{\tau} \in \mathcal{T} \mid \sum_{j=1}^{m} \tau_j \leq K\}$.

Regarding K, a practical bound is suggested in Reveliotis (2005) for a specific class of PNs (namely, process-resource nets with acyclic, quasi-live and strongly reversible process subnets). Furthermore, in Basile et al. (2012) a necessary and sufficient condition is provided for the K-diagnosability of bounded nets. For bounded net systems, there exists an integer $K_{\min}$ such that $K_{\min}$ FCVs are sufficient to *span* the reachability set $R(N, \boldsymbol{m}_0)$. Although in the worst case $K_{\min}$ may be equal to card $(R\,(N,\ \boldsymbol{m}_0)) - 1$, in many cases $K_{\min} \ll$ card $(R\,(N,\ \boldsymbol{m}_0)) - 1$. For its estimation in bounded and live systems the interested reader can refer to Sec. 3 in Basile et al. (2012). In control oriented problems it is sometimes possible to provide a meaningful pricing of K, if a cost (*e.g.*, in terms of the required task time) is associated with the firing of transitions. For example, in predictive control approaches at each decision step a limited forward horizon is scanned (in terms of number of firings (Lefebvre 2016) or total time). More in general, a limit can be posed on the cost of acceptable solutions.

Finally, notice that linear constraints such as generalized mutual exclusion constraints (GMECs) (Giua et al. 1992) on the marking variables can be seamlessly included in the formulation, if desired. More specifically, if the GMECs are in the form $\boldsymbol{L}\boldsymbol{m} \leq \boldsymbol{b}$, one has to add the following constraints:

$$\boldsymbol{L}\left(\boldsymbol{m}_0 + \boldsymbol{C}\sum_{k=1}^{i-1}\boldsymbol{e}^{(k)}\right) \leq \boldsymbol{b}$$

for $i = 1, \ldots, K$.

3.2 Problem 2: Reaching a set of target markings with or without via markings

It is often the case that the reachability target is a set of markings as opposed to a specific one, meaning that the objective is achieved when one of the markings of the set is reached. Suppose, *e.g.*, that the target set is defined as $M_t = \{\boldsymbol{m}_{t,1}, \boldsymbol{m}_{t,2}, \ldots, \boldsymbol{m}_{t,q}\}$. Then, we need to make the following substitution in the formulation:

$$\boldsymbol{m}_t = w_1\boldsymbol{m}_{t,1} + w_2\boldsymbol{m}_{t,2} + \ldots + w_q\boldsymbol{m}_{t,q},$$

where the coefficients w_l, $l = 1, \ldots, q$, are binary and sum up to one.

Note that the condition

$$\begin{bmatrix}\boldsymbol{\sigma}\\ 1\end{bmatrix} = \sum_{j=1}^{p}\alpha_j\boldsymbol{\tau}^{(j)}$$

should be updated to account for all elements of $\cup_{l=1}^{q}\mathcal{T}^{\mathrm{ms},l}$, where $\mathcal{T}^{\mathrm{ms},l}$ is the set of the minimal-support T-invariants of the complemented net associated to the target marking $\boldsymbol{m}_{t,l}$. Furthermore, a mutual exclusion contraint should also be imposed on the α_j coefficients associated to different sets of minimal-support T-invariants. It is therefore arguable at this point that going through the added effort of computing beforehand the minimal-support T-invariants of all the required complemented nets is actually beneficial, given that it ultimately introduces a further layer of complexity in the formulation. We suggest dropping altogether the condition based on the SCTs, on the grounds that this does not invalidate the formulation, given the reachability condition expressed by means of the state equation and the admissibility constraints, although it may impact in the solution efficiency.

Overall, the problem formulation is modified as follows:

$$
\begin{aligned}
&\text{minimize} \sum_{k=1}^{K}\sum_{j=1}^{m} e_j^{(k)} \\
&\text{subject to} \\
&\sum_{l=1}^{q} w_l\, \boldsymbol{m}_{t,l} = \boldsymbol{m}_0 + \boldsymbol{C}\sum_{k=1}^{K} \boldsymbol{e}^{(k)} \\
&\boldsymbol{m}_0 + \boldsymbol{C}\sum_{k=1}^{i-1} \boldsymbol{e}^{(k)} \geq \boldsymbol{Pre}\ \boldsymbol{e}^{(i)} && i = 1, \ldots, K \\
&\sum_{j=1}^{m} e_j^{(k)} \leq 1 && k = 1, \ldots, K \\
&\sum_{j=1}^{m} e_j^{(k)} \geq \sum_{j=1}^{m} e_j^{(k+1)} && k = 1, \ldots, K-1 \\
&e_j^{(k)} \in \{0, 1\} && j = 1, \ldots, m,\ k = 1, \ldots, K \\
&\sum_{l=1}^{q} w_l = 1 \\
&w_l \in \{0, 1\} && l = 1, \ldots, q
\end{aligned}
$$

3.3 Problem 3: Constraining the sequence to pass through intermediate markings

In navigational problems, the target of the reachability task may also involve the request to pass through some intermediate states before completing the trajectory, either in a given order or not. Again, in terms of SCTs this would require the precomputation of all SCTs from $\boldsymbol{m}_0$ to all way points, and again from there to $\boldsymbol{m}_t$ and suitably combining them to obtain the full FCVs. As explained in the previous subsection the complexity would increase even further if marking sets are used as opposed to single markings to define the way points or the target markings. We here introduce a formulation that avoids resorting to an explicit pre-computation of the SCTs, and is only based on the state equation and the admissibility conditions.

Assume first that a single target point $\boldsymbol{m}_t$ is specified, and it is requested that all trajectories pass through the waypoint $\boldsymbol{m}_w$. Then, the formulation of Problem 1 must be augmented as follows:

$$
\begin{aligned}
&\boldsymbol{m}_0 + \boldsymbol{C}\sum_{k=1}^{K} \beta_k \boldsymbol{e}^{(k)} = \boldsymbol{m}_w \\
&\beta_k \geq \beta_{k+1} && k = 1, \ldots, K-1 \\
&\beta_1 = 1 \\
&\beta_K = 0 \\
&\beta_k \in \{0, 1\} && k = 2, \ldots, K-1
\end{aligned}
$$

where the first equation is the state equation limited to the first part of the transition sequence, expressing the constraint that the trajectory should reach $\boldsymbol{m}_w$. The number of elements of $\boldsymbol{\beta}$ equal to 1 identifies the step at which the waypoint is actually reached. Notice that the first equation is actually a nonlinear constraint, as both β_k and $\boldsymbol{e}^{(k)}$ are variables. However, it can be easily linearized by introducing the additional binary variables $\delta_{jk} = \beta_k e_j^{(k)}$, which can be characterized as follows:

$$\begin{aligned}
&\delta_{jk} \leq \beta_k \\
&\delta_{jk} \leq e_j^{(k)} \\
&\delta_{jk} \geq \beta_k + e_j^{(k)} - 1 \\
&\delta_{jk} \in \{0, 1\}
\end{aligned}$$

for $j = 1, \ldots, m$, and $k = 1, \ldots, K$. Using the δ_{jk} variables, the state equation constraint becomes:

$$m_{0,i} + \sum_{k=1}^{K} \sum_{j=1}^{m} C_{ij} \delta_{jk} = m_{w,i}, \ i = 1, \ldots, n.$$

Overall, the full formulation becomes:

$$\begin{aligned}
&\text{minimize} \sum_{k=1}^{K} \sum_{j=1}^{m} e_j^{(k)} \\
&\text{subject to} \\
&\boldsymbol{m}_t = \boldsymbol{m}_0 + \boldsymbol{C} \sum_{k=1}^{K} \boldsymbol{e}^{(k)} \\
&\boldsymbol{m}_0 + \boldsymbol{C} \sum_{k=1}^{i-1} \boldsymbol{e}^{(k)} \geq \boldsymbol{Pre}\ \boldsymbol{e}^{(i)} && i = 1, \ldots, K \\
&\sum_{j=1}^{m} e_j^{(k)} \leq 1 && k = 1, \ldots, K \\
&\sum_{j=1}^{m} e_j^{(k)} \geq \sum_{j=1}^{m} e_j^{(k+1)} && k = 1, \ldots, K-1 \\
&e_j^{(k)} \in \{0, 1\} && j = 1, \ldots, m, \ k = 1, \ldots, K \\
&m_{0,i} + \sum_{k=1}^{K} \sum_{j=1}^{m} C_{ij} \delta_{jk} = m_{w,i} && i = 1, \ldots, n \\
&\beta_k \geq \beta_{k+1} && k = 1, \ldots, K-1 \\
&\beta_1 = 1 \\
&\beta_K = 0 \\
&\beta_k \in \{0, 1\} && k = 2, \ldots, K-1 \\
&\delta_{jk} \leq \beta_k && j = 1, \ldots, m, k = 1, \ldots, K
\end{aligned}$$

$$
\begin{array}{ll}
\delta_{jk} \le e_j^{(k)} & j = 1, \dots, m, k = 1, \dots, K \\
\delta_{jk} \ge \beta_k + e_j^{(k)} - 1 & j = 1, \dots, m, k = 1, \dots, K \\
\delta_{jk} \in \{0, 1\} & j = 1, \dots, m, k = 1, \dots, K
\end{array}
$$

If multiple waypoints are defined, the additional constraints (last 4 rows of the formulation) must be replicated for each of them (with personalized β_k and δ_{jk} variables). This forces the trajectory to include all waypoints. If, furthermore, we wish to impose a strict order among the waypoints we need to impose a simple constraint on the sum of the β_k coefficients associated to each waypoint, as this number coincides with the sequence step in which the waypoint is reached. Finally, if waypoints are not specific markings, but sets of markings, the same method used in Problem 2 for the target markings must be replicated for the waypoints as well.

4 The branching scheme

Using the previous ILP formulations as a cornerstone we can employ a B&B method to enumerate all admissible FCVs with not more than K firings of the original transitions that satisfy the reachability requirements. Two different versions of the branching method can be envisaged depending on whether we are interested only in the minimal among such admissible FCVs or in all of them. For example, a non-minimal admissible FCV could be obtained by summing an NCT to an admissible SCT (this occurs when a fireable sequence is temporarily interrupted at an intermediate point to follow a complete T-invariant, and then is resumed). If one is strictly interested in the shortest paths leading to a certain state, then the first version of the branching method is appropriate (interrupting the appropriate sequence to follow a T-invariant is just a waste of time). However, in certain timed control problems it may be necessary to spend some time in idle operations before undertaking a certain task. In this second case, non-minimal admissible FCVs might also be of interest.

The B&B method operates by partitioning the solution space into smaller regions that exclude all previously found solutions, and –if desired– all the solutions that are (element-wise) greater than or equal to those already found (non-minimal solutions). The following property defines the said partition.

Property 1 *Let $\Sigma \neq \emptyset$ be the (non-empty) set of all admissible FCVs satisfying the reachability conditions, and let $\overline{\Sigma} \subseteq \Sigma$ be a generic non-empty subset. Let also $\boldsymbol{\sigma}^\star \in \overline{\Sigma}$ be such that there does not exist $\boldsymbol{\sigma} \in \overline{\Sigma} \setminus \{\boldsymbol{\sigma}^\star\}$ such that $\sum_{j=1}^m \sigma_j < \sum_{j=1}^m \sigma_j^\star$. Then $\overline{\Sigma}$ can be partitioned as:*

$$\overline{\Sigma} = \{\boldsymbol{\sigma}^\star\} \cup \overline{\Sigma}^{(1)} \cup \ldots \cup \overline{\Sigma}^{(m)} \cup \overline{\Sigma}^{(m+1)}$$

where $\overline{\Sigma}^{(j)} = \{\boldsymbol{\sigma} \in \overline{\Sigma} \mid (\sigma_l \ge \sigma_l^\star, l = 1, \dots, j-1) \wedge (\sigma_j < \sigma_j^\star)\}$, $j = 1, \dots, m$, and $\overline{\Sigma}^{(m+1)} = \{\boldsymbol{\sigma} \in \overline{\Sigma} \mid (\sigma_j \ge \sigma_j^\star, j = 1, \dots, m) \wedge (\sum_{j=1}^m \sigma_j > \sum_{j=1}^m \sigma_j^\star)\}$.

Notice that:

- $\overline{\Sigma}^{(j)} \cap \overline{\Sigma}^{(l)} = \emptyset, j, l = 1, \dots, m+1, j \neq l.$
- None of the sets $\overline{\Sigma}^{(j)}$, $j = 1, \dots, m+1$ includes $\boldsymbol{\sigma}^\star$ (since either one element of $\boldsymbol{\sigma}$ or the sum of its elements are forced to be different) or a solution with a smaller sum of elements (by assumption).
- Only the set $\overline{\Sigma}^{(m+1)}$ can include solutions that are element-wise greater than or equal to $\boldsymbol{\sigma}^\star$ (non-minimal solutions).

In view of the previous partition, if an optimal solution to the ILP problem $\Pi_0 = (f_0, \mathcal{C}_0)$ is found, corresponding to a firing vector $\boldsymbol{\sigma}^\star$, further minimal solutions can be sought for by addressing the following modified versions of the same problem:
$\Pi_1 = (f_0, \mathcal{C}_1)$, where $\mathcal{C}_1 = \mathcal{C}_0 \wedge (\sigma_1 < \sigma_1^\star)$,
$\Pi_2 = (f_0, \mathcal{C}_2)$, where $\mathcal{C}_2 = \mathcal{C}_0 \wedge (\sigma_1 \geq \sigma_1^\star) \wedge (\sigma_2 < \sigma_2^\star)$,
…
$\Pi_m = (f_0, \mathcal{C}_m)$, where $\mathcal{C}_m = \mathcal{C}_0 \wedge (\sigma_j \geq \sigma_j^\star, j = 1, \ldots, m-1) \wedge (\sigma_m < \sigma_m^\star)$.
If non-minimal solutions are of interest as well, a further problem must be added:
$\Pi_{m+1} = (f_0, \mathcal{C}_{m+1})$, where $\mathcal{C}_{m+1} = \mathcal{C}_0 \wedge (\sigma_j \geq \sigma_j^\star, j = 1, \ldots, m) \wedge (\sum_{j=1}^m \sigma_j > \sum_{j=1}^m \sigma_j^\star)$.

The partitioning procedure of Property 1 is applied again for each of the generated sub-problems that admit a solution, thereby configuring a tree of sub-problems (also called nodes) stemming from Π_0, each one associated with a specific subset of Σ. Notice that all nodes including a constraint of the type $\sigma_j < \sigma_j^\star$ can be immediately classified as infeasible if $\sigma_j^\star = 0$. When the tree expansion is completed (*i.e.* all leaves of the branching tree correspond to infeasible sub-problems), the algorithm ends and the solutions found at the non-leaf nodes constitute either the full solution set $\Sigma^\star$ (limited to sequences of length not greater than K steps) or its subset containing the minimal solutions, depending on which version of the branching rule has been applied.

The branching procedure can be summarized as follows (Step 4 is to be included only if non-minimal solutions are of interest):

```
Input:   $\Pi_0$ (basic ILP problem formulation)
Output:  $\Sigma^\star$ (solution set)
Step 0)  Initialize the list[1] of open problems, $\Lambda = \langle \Pi_0 \rangle$.
         Initialize the solution set, $\Sigma^\star = \emptyset$.
Step 1)  If $\Lambda = \langle\rangle$ return $\Sigma^\star$ and exit, else $\Pi = (f, \mathcal{C}) = \texttt{pop}(\Lambda)$.
Step 2)  $\boldsymbol{\sigma}^\star = \texttt{solve}(\Pi)$[2]. If $\boldsymbol{\sigma}^\star = \emptyset$, go to Step 1, else $\Sigma^\star = \Sigma^\star \cup \{\boldsymbol{\sigma}^\star\}$.
Step 3)  For $j = 1$ to $m$ do $\Lambda = \langle \Lambda, (f, \mathcal{C}') \rangle$, with $\mathcal{C}' = \mathcal{C} \wedge (\sigma_k \geq \sigma_k^\star, k = 1, \ldots, j-1) \wedge$
         $(\sigma_j < \sigma_j^\star)$.
Step 4)  [optional] $\Lambda = \langle \Lambda, (f, \mathcal{C}') \rangle$, with $\mathcal{C}' = \mathcal{C} \wedge (\sigma_j \geq \sigma_j^\star, j = 1, \ldots, m) \wedge (\sum_{j=1}^m \sigma_j >$
         $\sum_{j=1}^m \sigma_j^\star)$.
Step 5)  Go to Step 1.
```

5 Examples

5.1 Example 1: Problem 1 for minimal solutions only

Consider the PN depicted in Fig. 1, with $n = 9$ places and $m = 10$ transitions (Kostin 2008), for which we want to solve the reachability problem from $\boldsymbol{m}_0 = [2\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]^T$ to $\boldsymbol{m}_t = [2\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1]^T$. We first calculate the minimal-support T-invariants of the complemented net, obtaining as in Kostin (2008), $\mathcal{T}^{\text{ms}} = \{\boldsymbol{\tau}^{(1)}, \boldsymbol{\tau}^{(2)}, \boldsymbol{\tau}^{(3)}\}$, where:
$\boldsymbol{\tau}^{(1)} = [0\ 0\ 1\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 0]^T$,

---

[1] A list Λ is an ordered set of elements $\Lambda = \langle \lambda_1, \lambda_2, \ldots, \lambda_k \rangle$. The `pop` function extracts the first element of a list, reducing it to $\Lambda = \langle \lambda_2, \ldots, \lambda_k \rangle$. Two lists $\Lambda = \langle \lambda_1, \ldots, \lambda_k \rangle$ and $M = \langle \mu_1, \ldots, \mu_j \rangle$ can be appended as follows: $\langle \Lambda, M \rangle = \langle \lambda_1, \ldots, \lambda_k, \mu_1, \ldots, \mu_j \rangle$. Finally, an empty list is denoted as $\langle\rangle$.

[2] `solve`(Π) indicates the solution of problem Π to optimality by means of a MIP solver. If no solution exists we assume that the procedure returns an empty set.

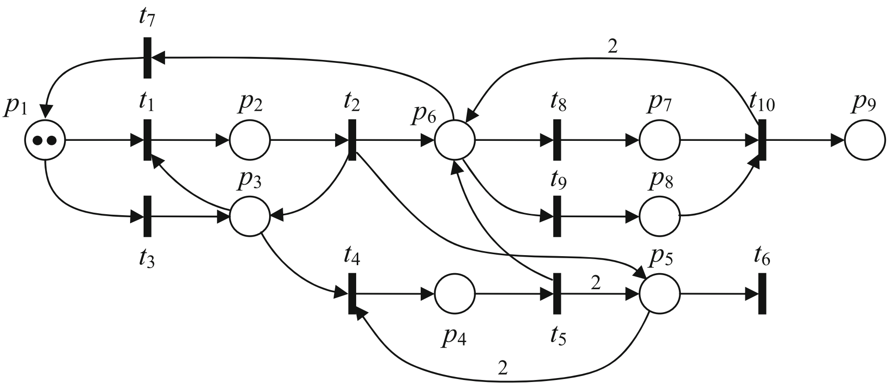

Fig. 1 PN of Example 1

$\boldsymbol{\tau}^{(2)} = [1\ 1\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0]^T$,
$\boldsymbol{\tau}^{(3)} = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1]^T$.
The first two are NCTs (their last element is 0) and correspond to the T-invariants of the original PN. The last one is a SCT, corresponding to a FCV $[0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1]^T$ satisfying the required marking variation. As the reader can immediately verify, however, there does not exist an enabled transition sequence compatible with such FCV. This does not necessarily imply that the requested reachability problem cannot be solved.

Indeed, the ILP corresponding to the 1^{st} formulation (with $K = 25$) yields a feasible solution corresponding to the shortest transition sequence achieving the wanted result, *i.e.* ($t_3\ t_1\ t_2\ t_7\ t_1\ t_2\ t_4\ t_9\ t_5\ t_8\ t_{10}\ t_7\ t_7\ t_6\ t_6$), corresponding to the SCT with non-minimal-support $\boldsymbol{\tau}^{(4)} = [2\ 2\ 1\ 1\ 1\ 2\ 3\ 1\ 1\ 1\ 1]^T$, which equals $\boldsymbol{\tau}^{(1)} + 2\boldsymbol{\tau}^{(2)} + \boldsymbol{\tau}^{(3)}$. In other words, to reach $\boldsymbol{m}_t$ one has still to follow the FCV $[0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1]^T$, but in order to enable the sequence must *borrow* the necessary tokens through suitable executions of the two T-invariants of the original PN. Notice that there may be alternative legal firing sequences compatible with $\boldsymbol{\tau}^{(4)}$ (indeed, the one documented in Kostin 2008 is a permutation of ours).

Branching the problem (according to the first version of the method) yields 10 children nodes, with the following additional sets of constraints, respectively:
node 1: $\sigma_1 < 2$
node 2: $\sigma_1 \geq 2$ and $\sigma_2 < 2$
node 3: $\sigma_1 \geq 2$, $\sigma_2 \geq 2$, and $\sigma_3 < 1$
etc.
None of these sub-problems provides feasible solutions, suggesting that only non-minimal solutions obtained by combining $\boldsymbol{\tau}^{(1)}$, $\boldsymbol{\tau}^{(2)}$, and $\boldsymbol{\tau}^{(3)}$ with higher coefficients may be found.

5.2 Example 2: Problem 1 for minimal and non-minimal solutions

Consider now the PN shown in Fig. 2, with $n = 9$ places and $m = 13$ transitions (Kostin 2008), for which we want to solve the reachability problem from $\boldsymbol{m}_0 = [1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]^T$ to $\boldsymbol{m}_t = [1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1]^T$ with transition firing sequences not longer than $K = 25$ steps. This time the set of minimal-support T-invariants of the complemented net $\mathcal{T}^{\text{ms}} = \{\boldsymbol{\tau}^{(1)}, \ldots, \boldsymbol{\tau}^{(7)}\}$ is much larger, with
$\boldsymbol{\tau}^{(1)} = [1\ 1\ 1\ 2\ 2\ 3\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]^T$,
$\boldsymbol{\tau}^{(2)} = [2\ 2\ 2\ 1\ 1\ 0\ 3\ 0\ 0\ 0\ 0\ 0\ 0\ 0]^T$,

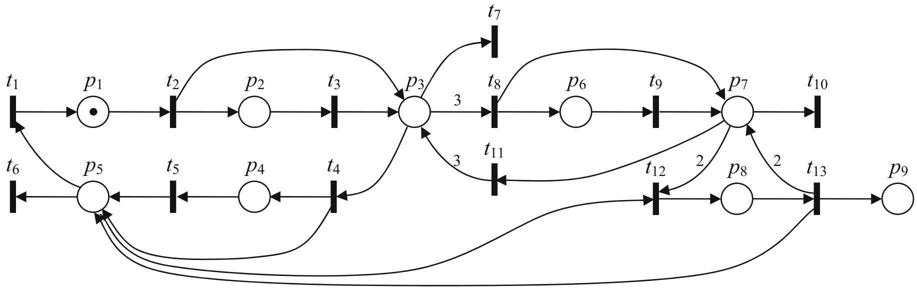

Fig. 2 PN of Example 2

$\boldsymbol{\tau}^{(3)} = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1\ 0\ 0\ 0]^T$,
$\boldsymbol{\tau}^{(4)} = [0\ 0\ 0\ 0\ 0\ 0\ 3\ 1\ 1\ 0\ 2\ 0\ 0\ 0]^T$,
$\boldsymbol{\tau}^{(5)} = [0\ 0\ 0\ 3\ 3\ 6\ 0\ 1\ 1\ 0\ 2\ 0\ 0\ 0]^T$,
$\boldsymbol{\tau}^{(6)} = [2\ 2\ 2\ 1\ 1\ 0\ 0\ 1\ 1\ 2\ 0\ 0\ 0\ 0]^T$,
$\boldsymbol{\tau}^{(7)} = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1]^T$,
the last of which being an SCT, while the first six are NCTs.

As in the previous example, the FCV $[0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1]^T$ corresponding to the only minimal-support SCT $\boldsymbol{\tau}^{(7)}$ turns out to be inadmissible (there does not exist a fireable transition sequence compatible with it). By applying the method in its second version (that is, including also non-minimal solutions) one obtains the branching tree represented in Fig. 3, where only the feasible nodes are reported for the sake of simplicity. More in detail, the minimal solution obtained with the basic formulation (node 1) is the FCV $\boldsymbol{\sigma}^{(1)} =$

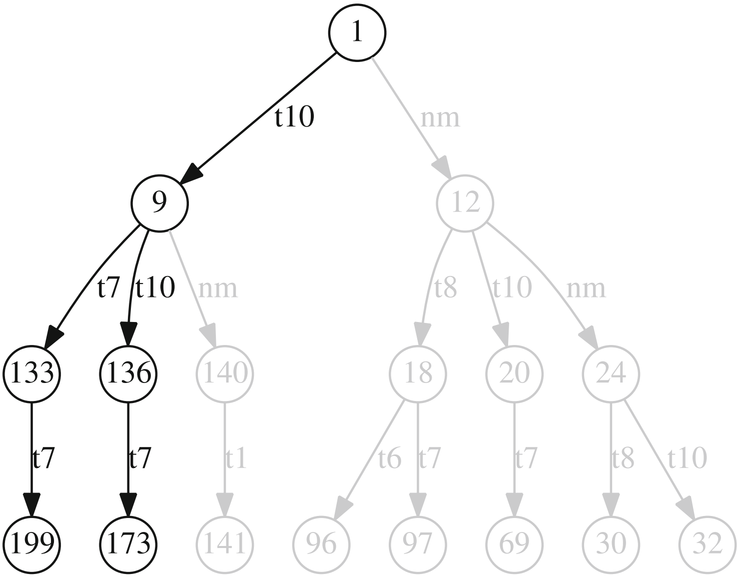

Fig. 3 Branching tree for Example 2: only feasible nodes are represented and arcs are labeled with the transition inducing the branch (label “nm” indicates the optional branching rule for non-minimal solutions). Nodes corresponding to non-minimal solutions are colored in gray

$[2\ 2\ 2\ 1\ 1\ 0\ 0\ 1\ 1\ 2\ 0\ 1\ 1]^T$, corresponding to an SCT obtained by summing together $\boldsymbol{\tau}^{(6)}$ and $\boldsymbol{\tau}^{(7)}$ (the combination of an NCT with an SCT is still an SCT). $\boldsymbol{\sigma}^{(1)}$ is the admissible FCV with the fewest possible transition firings. The branching process generates $m+1 = 14$ nodes, resulting in a full partition of the solution space (only solutions such that $\boldsymbol{\sigma} \leq \boldsymbol{\sigma}^{(1)}$ are excluded from further consideration, since there is no smaller admissible solution than $\boldsymbol{\sigma}^{(1)}$). The first node (node 2) is obtained by adding the constraint $\sigma_1 < 2$ to the basic formulation, but is not represented as it doesn't produce admissible solutions. The second node (node 3) is obtained by adding the constraints $\sigma_1 \geq 2$ and $\sigma_2 < 2$ to the original problem. Again this node is not represented in the picture, as it does not provide further solutions. Notice that three children nodes (precisely, those corresponding to the null elements of $\boldsymbol{\sigma}^{(1)}$) can be safely omitted, as the added constraints are clearly unfeasible. Therefore, of the 13 potential nodes, only 10 are actually analyzed (nodes 2 to 11), but just one is represented, as the other ones do not produce any solution. The successful node (node 9) is generated by adding the constraints: $\sigma_1 \geq 2$, $\sigma_2 \geq 2$, $\sigma_3 \geq 1$, $\sigma_4 \geq 1$, $\sigma_5 \geq 1$, $\sigma_8 \geq 1$, $\sigma_9 \geq 1$, and $\sigma_{10} < 2$. The solution found in node 9 is $\boldsymbol{\sigma}^{(2)} = [2\ 2\ 2\ 1\ 1\ 0\ 3\ 1\ 1\ 1\ 1\ 1\ 1]^T$, which indeed satisfies all these conditions. The last node generated from the branching of node 1 is tasked with finding a solution $\boldsymbol{\sigma} \geq \boldsymbol{\sigma}^{(1)}$, but such that $\boldsymbol{\sigma} \neq \boldsymbol{\sigma}^{(1)}$ (node 12). This node also produces an admissible FCV, namely $\boldsymbol{\sigma}^{(3)}$, that is indeed such that $\boldsymbol{\sigma}^{(3)} \geq \boldsymbol{\sigma}^{(1)}$. The algorithm continues by further branching nodes 9 and 12. Overall, 238 nodes are explored, 17 of which correspond to feasible solutions (*i.e.*, admissible SCTs, with $\bar{k} \leq K$ transition firings), as reported in Table 1.

For each of these FCVs there exists at least one fireable transition sequence of the PN (the algorithm returns such sequence as well), taking it from $\boldsymbol{m}_0$ to $\boldsymbol{m}_t$. For example, for $\boldsymbol{\sigma}^{(1)}$ one such sequence is ($t_2\ t_4\ t_1\ t_2\ t_3\ t_3\ t_8\ t_9\ t_5\ t_{12}\ t_{13}\ t_{10}\ t_{10}\ t_1$). Notice that only $\boldsymbol{\sigma}^{(1)}$, $\boldsymbol{\sigma}^{(2)}$, $\boldsymbol{\sigma}^{(12)}$, $\boldsymbol{\sigma}^{(13)}$, $\boldsymbol{\sigma}^{(16)}$, $\boldsymbol{\sigma}^{(17)}$ are minimal solutions. All the other solutions are not minimal. Some of

Table 1 Example 2: admissible FCVs

node	admissible FCVs	$\bar{k}$	min.
1	$\boldsymbol{\sigma}^{(1)} = [2\ 2\ 2\ 1\ 1\ 0\ 0\ 1\ 1\ 2\ 0\ 1\ 1]^T$	14	Y
9	$\boldsymbol{\sigma}^{(2)} = [2\ 2\ 2\ 1\ 1\ 0\ 3\ 1\ 1\ 1\ 1\ 1\ 1]^T$	17	Y
12	$\boldsymbol{\sigma}^{(3)} = [2\ 2\ 2\ 1\ 1\ 0\ 0\ 2\ 2\ 3\ 1\ 1\ 1]^T$	18	N
18	$\boldsymbol{\sigma}^{(4)} = [3\ 3\ 3\ 2\ 2\ 1\ 1\ 1\ 1\ 2\ 0\ 1\ 1]^T$	22	N
20	$\boldsymbol{\sigma}^{(5)} = [2\ 2\ 2\ 1\ 1\ 0\ 3\ 2\ 2\ 2\ 2\ 1\ 1]^T$	21	N
24	$\boldsymbol{\sigma}^{(6)} = [2\ 2\ 2\ 1\ 1\ 0\ 0\ 3\ 3\ 4\ 2\ 1\ 1]^T$	22	N
30	$\boldsymbol{\sigma}^{(7)} = [3\ 3\ 3\ 2\ 2\ 1\ 1\ 2\ 2\ 3\ 1\ 1\ 1]^T$	25	N
32	$\boldsymbol{\sigma}^{(8)} = [2\ 2\ 2\ 1\ 1\ 0\ 3\ 3\ 3\ 3\ 3\ 1\ 1]^T$	25	N
69	$\boldsymbol{\sigma}^{(9)} = [2\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 2\ 1\ 1]^T$	24	N
96	$\boldsymbol{\sigma}^{(10)} = [4\ 4\ 4\ 2\ 2\ 0\ 3\ 1\ 1\ 2\ 0\ 1\ 1]^T$	25	N
97	$\boldsymbol{\sigma}^{(11)} = [3\ 3\ 3\ 3\ 3\ 3\ 0\ 1\ 1\ 2\ 0\ 1\ 1]^T$	24	N
133	$\boldsymbol{\sigma}^{(12)} = [2\ 2\ 2\ 2\ 2\ 2\ 2\ 1\ 1\ 1\ 1\ 1\ 1]^T$	20	Y
136	$\boldsymbol{\sigma}^{(13)} = [2\ 2\ 2\ 1\ 1\ 0\ 6\ 1\ 1\ 0\ 2\ 1\ 1]^T$	20	Y
140	$\boldsymbol{\sigma}^{(14)} = [3\ 3\ 3\ 2\ 2\ 1\ 4\ 1\ 1\ 1\ 1\ 1\ 1]^T$	24	N
141	$\boldsymbol{\sigma}^{(15)} = [2\ 2\ 2\ 1\ 1\ 0\ 6\ 2\ 2\ 1\ 3\ 1\ 1]^T$	24	N
173	$\boldsymbol{\sigma}^{(16)} = [2\ 2\ 2\ 2\ 2\ 2\ 5\ 1\ 1\ 0\ 2\ 1\ 1]^T$	23	Y
199	$\boldsymbol{\sigma}^{(17)} = [2\ 2\ 2\ 3\ 3\ 4\ 1\ 1\ 1\ 1\ 1\ 1\ 1]^T$	23	Y

these non-minimal solutions are even shorter than some given minimal ones in terms of the sequence length $\overline{k}$. Notice that all solutions correspond to suitable linear combinations of the SCT $\boldsymbol{\tau}^{(7)}$ with the 6 NCTs. A few of these solutions are also documented in Kostin (2008), namely $\boldsymbol{\sigma}^{(1)}$, $\boldsymbol{\sigma}^{(4)}$, and $\boldsymbol{\sigma}^{(9)}$.

5.3 Example 3: Problems 2 and 3

Consider again the PN of Fig. 1, with $\boldsymbol{m}_0$ and $\boldsymbol{m}_t$ defined as before, but this time we require that the trajectory should pass also by an assigned waypoint. Assume first that the waypoint is $\boldsymbol{m}_{w1} = [0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 0]^T$. Applying the 3^{rd} problem formulation to this case we obtain the following list of admissible FCVs (with no more than $K = 25$ steps):

- $\boldsymbol{\sigma}^{(1)} = [3\ 3\ 1\ 1\ 1\ 3\ 4\ 1\ 1\ 1]$ (19 steps)
- $\boldsymbol{\sigma}^{(2)} = [4\ 4\ 1\ 1\ 1\ 4\ 5\ 1\ 1\ 1]$ (23 steps)
- $\boldsymbol{\sigma}^{(3)} = [3\ 3\ 2\ 2\ 2\ 3\ 5\ 1\ 1\ 1]$ (23 steps)

For example, following the sequence ($t_3\ t_1\ t_2\ t_6\ t_7\ t_1\ t_2\ t_7\ t_1\ t_2\ t_4\ t_9\ t_5\ t_8\ t_{10}\ t_7\ t_7\ t_6\ t_6$), corresponding to $\boldsymbol{\sigma}^{(1)}$, $\boldsymbol{m}_{w1}$ is reached after 4 steps and $\boldsymbol{m}_t$ after 19. This sequence is longer than that found previously in Section 5.1, due to the additional constraint introduced by the waypoint. As explained in Section 5.1 the found solution must be a non-minimal FCV linearly combining the basic FCV leading to $\boldsymbol{m}_t$ (see $\boldsymbol{\tau}^{(3)}$ in Section 5.1) with the two T-invariants ($\boldsymbol{\tau}^{(1)}$ and $\boldsymbol{\tau}^{(2)}$), with higher coefficients than the unconstrained FCV. It can be easily verified that this indeed holds. Notice that the other two FCVs of length 23 are obtained by further extending the FCV by either T-invariant.

Using a different waypoint, $\boldsymbol{m}_{w2} = [0\ 0\ 2\ 0\ 2\ 0\ 0\ 0\ 0]^T$, yields the following FCVs instead:

- $\boldsymbol{\sigma}^{(4)} = [2\ 2\ 2\ 2\ 2\ 2\ 4\ 1\ 1\ 1]$ (19 steps)
- $\boldsymbol{\sigma}^{(5)} = [2\ 2\ 3\ 3\ 3\ 2\ 5\ 1\ 1\ 1]$ (23 steps)
- $\boldsymbol{\sigma}^{(6)} = [3\ 3\ 2\ 2\ 2\ 3\ 5\ 1\ 1\ 1]$ (23 steps)

Notice that the shortest FCV is different from the one found for $\boldsymbol{m}_{w1}$. However, it holds that $\boldsymbol{\sigma}^{(3)} = \boldsymbol{\sigma}^{(6)}$.

Suppose now that we allow *either* waypoint to be reached in the trajectory, using the same method employed for multiple target markings in formulation 2. Not surprisingly, the result of the enumeration process is the union of the FCVs found for each waypoint.

If, on the other hand, we impose that *both* waypoints should be reached, in whatever order (see formulation 3), the number of possible solutions reduces to just one:

- $\boldsymbol{\sigma}^{(7)} = [3\ 3\ 2\ 2\ 2\ 3\ 5\ 1\ 1\ 1]$ (23 steps)

For example, it takes 4 steps for the sequence ($t_3\ t_1\ t_2\ t_6\ t_7\ t_1\ t_2\ t_7\ t_1\ t_2\ t_7\ t_3\ t_4\ t_5\ t_4\ t_8\ t_5\ t_9\ t_6\ t_{10}\ t_7\ t_6\ t_7$) to reach $\boldsymbol{m}_{w1}$ and 12 to arrive to $\boldsymbol{m}_{w2}$, before finishing in $\boldsymbol{m}_t$ in 23 steps.

Clearly, the same solution is found if the order $\boldsymbol{m}_{w1} - \boldsymbol{m}_{w2}$ is enforced, whereas the reverse order does not yield any solution (to find a solution, K must be increased to 27).

6 Conclusion

A method was proposed to solve in practice the problem of finding (and enumerating) the transition paths leading from the current state to one or a set of target states. The proposed approach exploits the mathematical representation of PNs and employs an ILP problem as the core element of a B&B approach to enumerate such transition paths. This formulation is particularly convenient since commercial optimization tools that are available off-the-shelf can be employed for its solution. In Watanabe et al. (1989) it is proved that finding legal firing sequences in general PNs is NP-complete in general. However, Integer Linear Programming (ILP) problems are standard optimization problems that can be efficiently solved by using off-the-shelf software, such as CPLEX and Fico-Xpress. Nowadays, such efficient commercial tools solve relatively large problems requiring a computational time that is compatible with several real-world applications, such as the industrial ones. Moreover, they are well suited for PN models with a high level of parallelism (Basile et al. 2018).

The approach specifies an upper bound for the number of events that can fire before a target state is reached. Given an integer K, we provide a set of conditions that need to be satisfied for an admissible transition path of length not greater than K to lead from the current state to the target one. These conditions are based on the state equation, transition admissibility, and some recently developed algebraic concepts related to firing count vectors.

This practical approach allows to verify if the target can be reached within a specified maximum time delay. Indeed, if the maximum interleaving time between two firings is given, the number of firings can be directly related to the time duration of the corresponding process, so that posing a bound on the arrival time is tantamount to limit the sequence length to K.

The basic formulation is then extended to deal with reachability requirements involving waypoints (possibly more than one, either in a given order or not) and sets of markings (as opposed to single markings). In this way, complex navigational problems can be addressed without having to perform a full reachability analysis (only sequences compatible with the found FCVs are explored).

Author Contributions All the authors contributed to the manuscript.

Funding Open access funding provided by Università degli Studi di Salerno within the CRUI-CARE Agreement. There was no funding for this paper.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Competing Interests The authors declare no competing interests.

References

Basile F, Chiacchio P, Tommasi GD (2012) On K-diagnosability of Petri nets via integer linear programming. Automatica 48(9):2047–2058

Basile F, Cordone R, Piroddi L (2015) A branch and bound approach for the design of decentralized supervisors in Petri net models. Automatica 52:322–333

Basile F, Cordone R, Piroddi L (2022) Supervisory control of timed discrete event systems with logical and timed specifications. IEEE Trans Autom Control 67(6):2800–2815

Basile F, Boussif A, De Tommasi G, Ghazel M, Sterle C (2018) Efficient diagnosability assessment via ILP optimization: a railway benchmark. In: 23rd IEEE international conference on emerging technologies and factory automation. Torino, Italy, pp 441–448

Chouchane A, Ghazel M, Boussif A (2023) K-diagnosability analysis of bounded and unbounded petri nets using linear optimization. Automatica 147:110689

Colom JM, Silva M (1991) Convex geometry and semiflows in P/T nets. A comparative study of algorithms for computation of minimal p-semiflows. In: Rozenberg G (ed) Advances in Petri nets 1990, pp 79–112. Springer, Berlin, Heidelberg

Cong X, Fanti M, Mangini A, Li Z (2018) Decentralized diagnosis by Petri nets and integer linear programming. IEEE Trans Syst Man Cybern Syst 48(10):1689–1700

Cordone R, Piroddi L (2013) Parsimonious monitor control of Petri net models of FMS. IEEE Trans Syst Man Cybern Syst 43(1):215–221

Cordone R, Nazeem A, Piroddi L, Reveliotis S (2013) Designing optimal deadlock avoidance policies for sequential resource allocation systems through classification theory: existence results and customized algorithms. IEEE Trans Autom Control 58(11):2772–2787

Giua A, DiCesare F, Silva M (1992) Generalized mutual exclusion constraints on nets with uncontrollable transitions. 1992 IEEE int. conf. on systems, man, and cybernetics (Chicago, Illinois), pp 974–979

Kostin AE (2006) A reachability algorithm for general Petri nets based on transition invariants. In: Královič R, Urzyczyn P (eds) MFCS 2006. LNCS 4162, pp 608–621. Springer, Berlin, Heidelberg, Germany

Kostin AE (2008) Using transition invariants for reachability analysis of Petri nets. In: Kordic V (ed) Petri net, theory and applications, pp 435–458. I-Tech Education and Publishing, Vienna, Austria. Chap. 19

Kostin AE (2003) Reachability analysis in T-invariant-less Petri nets. IEEE Trans Autom Control 48(6):1019–1024

Lefebvre D (2016) Approaching minimal time control sequences for timed Petri nets. IEEE Trans Autom Sci Eng 13(2):1215–1221

Martínez J, Silva M (1982) A simple and fast algorithm to obtain all invariants of a generalised Petri net. In: Girault C, Reisig W (eds) Application and theory of Petri nets. Springer, Berlin, Heidelberg, pp 301–310

Mayr EW (1984) An algorithm for the general Petri net reachability problem. SIAM J Comput 13(3):441–459

Memmi G, Roucairol G (1980) Linear algebra in net theory. In: Proceedings of the advanced course on general net theory of processes and systems: net theory and applications, pp 213–223. Springer, London, UK, UK

Murata T (1989) Petri nets: properties, analysis and applications. Proc IEEE 77(4):541–580

Qi L, Su Y, Zhou M, Abusorrah A (2023) A state-equation-based backward approach to a legal firing sequence existence problem in petri nets. IEEE Trans Syst Man Cybern Syst 53(8):4968–4979

Reveliotis S (2005) A linear characterization of the Petri net reachability space corresponding to bounded-length fireable transition sequences and its implications for the structural analysis of process-resource nets with acyclic, quasi-live and strongly reversible process subnets. In: 44th IEEE conference on decision and control and European control conference, Seville, Spain, pp 2113–2118

Schmidt K (1999) Stubborn sets for standard properties. In: Donatelli S, Kleijn J (eds) Application and theory of Petri nets 1999. Springer, Berlin, Heidelberg, pp 46–65

Silva M, Teruel E, Colom JM (1998) In: Reisig W, Rozenberg G (eds) Linear algebraic and linear programming techniques for the analysis of place/transition net systems, pp 309–373. Springer, Berlin, Heidelberg

Su Y, Qi L, Zhou M (2023) A backward algorithm to determine the existence of legal firing sequences in ordinary petri nets. IEEE Robot Autom Lett 8(6):3190–3197

Watanabe T, Mizobata Y, Onaga K (1989) Legal firing sequence and related problems of petri nets. In: Proceedings of the third international workshop on Petri nets and performance models, PNPM89, pp 277–286

Roberto Cordone was born in Milan, Italy, in 1969. He received the “Laurea” degree in Electronic Engineering and the Ph.D. degree in Computer Science and Control Theory from the Politecnico di Milano, Milano, Italy, in 1996 and 2000, respectively. From 2002 to 2017 he was an Assistant Professor with the Universitá degli Studi di Milano. Since 2017 he is an Associate Professor with the Universitá degli Studi di Milano, where he holds courses on heuristic algorithms and decisions methods and models. His research interests include operations research and algorithm design and analysis.

Francesco Basile received the Laurea degree cum laude in electronic engineering and the Ph.D. degree in electronic and computer engineering from the University of Naples, Naples, in 1995 and 1999, respectively. In 1999, he was a Visiting Researcher with the Departamento de Ingenieria Informatica y Systemas, University of Zaragoza, Zaragoza, Spain, for six months. He is currently Full Professor of Automatic Control with the Dipartimento di Ingegneria dell’informazione ed elettrica e matematica applicata, Università di Salerno, Fisciano, Italy. He has published over 130 papers on international journals and conferences. His current research interests include modeling and control of discrete event systems, automated manufacturing, and robotics. Prof. Basile has been Associate Editor of the International Journal of Robotics and Automation, IEEE Transactions on Control Systems Technology, IEEE Transactions on Automation Science and Engineering and IEEE Control Systems Letters. He has been member of IEEE Control System Society Conference Editorial Board. He is Associate Editor of IEEE Transactions on Automatic Control. He has been General Chair of 14th International Workshop on Discrete Event Systems (WODES 2018).

Luigi Piroddi was born in London, U.K., in 1966. He received his laurea degree in Electrical Engineering and the Ph.D. degree in Computer Science and Control Theory from the Politecnico di Milano, Milano, Italy, in 1990 and 1995, respectively. Between 1994 and 1999, he was a Professor of Fundamentals of Systems and Control with the Universitá degli Studi di Bergamo, Bergamo, Italy. From 1999 to 2004, he was an Assistant Professor with the Politecnico di Milano. From 2004 to 2015 he has been an Associate Professor, and from 2016 he is Full Professor with the same institution, where he holds various courses in the systems and control area. His research interests include nonlinear model identification, Petri nets, modeling, and control of manufacturing processes.