Basic Linear Algebraic Techniques for Place/Transition Nets

Jörg Desel

Institut für Angewandte Informatik und Formale Beschreibungsverfahren
Universität Karlsruhe, Germany
desel@aifb.uni-karlsruhe.de

Abstract. Linear algebraic techniques for place/transition nets are surveyed. In particular, place and transition invariant vectors and their application to verification, proof and analysis of behavioral properties of marked Petri nets are presented. The considered properties are the non-reachability of a marking and conditions that hold true for all reachable markings. In addition, it is shown how the rank of the incidence matrix implies sufficient criteria and necessary criteria for liveness of bounded marked Petri nets.

1 Introduction

This contribution surveys the state of the art in linear algebraic techniques for Petri nets. More precisely, we show how the matrix representation of a Petri net together with appropriate equation systems or systems of inequalities can be exploited for gaining or proving properties of the net's behavior.

The main motivation for the linear algebraic approach is the behavioral complexity of Petri nets. In general, the set of reachable markings of an initially marked net can explode both with respect to the size of the net and the number of initial tokens, and it even can be infinite. Hence, in practice, the explicit construction of all reachable markings is not feasible. This implies that efficient analysis of behavioral properties can not be based on models of behavior such as marking graphs or sets of all occurrence sequences or even process nets. We rather have to stick to the structure of the considered net, together with its initial token distribution. Linear algebraic techniques exploit only those structural data. They yield at least sufficient *or* necessary conditions for properties of the behavior of a marked net. These conditions can be verified by help of efficient algorithms for linear algebraic techniques. The complexity of such algorithms heavily depend on the domain under consideration. For example, solutions of an equation system in the rational numbers are simpler to achieve than solutions in the integers or in the natural numbers. A trade-off frequently occurs between the complexity of an algorithm and the expressive power of its result.

Combining linear algebraic techniques and Petri nets is nearly as old as Petri net theory. The first relevant contributions appeared already in the mid 70ies. We collect both the old concepts and new results in a uniform and, hopefully, readable way. We do *not* consider arc weights, capacity restrictions, time annotations or other extensions. Most presented results can canonically be extended to these generalizations. We also do not consider specific properties of restricted classes of Petri nets such as free-choice Petri nets.

The applicability of equation systems to Petri nets is based on the duality of states and state changes that was emphasized by C. A. Petri [Petr73, Petr82]. In Petri nets, places, being local units of states, and transitions, being atomic units of state changes, are represented on the same level. This duality is reflected by linear methods: In the matrix representation of a net, places and transitions are represented by rows and columns, respectively. Linear algebra relates properties of a matrix to properties of its transposed matrix. This applies in particular to solubility criteria for the respective equation systems. Translated to Petri nets, these criteria yield relations between the places and the transitions of a net that are used in techniques described in this contribution.

All linear algebraic techniques for Petri nets are based on the observation that the occurrence of a transition always causes the same relative change of the marking of places: it decreases the number of tokens on each place which is in its pre-set but not in its post-set by one, and it increases the number of tokens on each place which is in its post-set but not in its pre-set by one. The markings of all other places remain unchanged. There exist modified occurrence rules for Petri nets that allow varying relative changes, too. For example, *preemptive arcs* allow to empty a place, no matter what its number of tokens was before. The more general *self-modifying nets* allow variable effects of a transition to a place, depending on the actual marking of some other place. These nets allow to copy the marking of some place to another place, similar to an assignment statement in a programming language. We will not consider such extensions in this paper.

As a general prerequisite for matrix and vector notations, we assume a finite set of places $S = \{s_1, \ldots, s_n\}$ and a finite set of transitions $T = \{t_1, \ldots, t_l\}$, that are indexed by the numbers $1, \ldots, n$ ($1, \ldots, l$, respectively). A marking associates to each place its actual number of tokens. Hence, every marking is a mapping $m: S \rightarrow \mathbb{N}$, where $\mathbb{N} = \{0, 1, 2, \ldots\}$ is the set of natural numbers (non-negative integers). A marking m can be represented by a vector $\mathbf{m} \in \mathbb{N}^n$ such that the i-th component of the vector $\mathbf{m}$ is the value $m(s_i)$. The constant relative change caused by a transition can be described by a vector, too. Its i-th component represents the effect of the transition's occurrence on the place s_i.

Figure 1 shows a vending-machine, as introduced in [DeRe98]. Its places are `s1`, `s2`, ..., `s6` and its transitions are `t1`, `t2`, ..., `t5`. The vector representation $\mathbf{m}_0$ of the initial marking and the vector representation $\mathbf{t2}$ of the transition `t2` read as

$$\mathbf{m}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{t2} = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}.$$

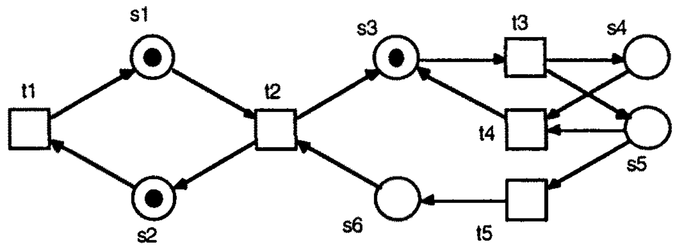


Fig. 1. A marked Petri net

A transition occurrence changes the number of tokens on a place at most by one. Hence, the vectors associated to the transitions have components of the set $\{-1, 0, 1\}$. Generalizations to arbitrary integers are not difficult and they are studied in the literature in depth (when nets with *weighted arcs* are considered).

If the occurrence of a transition t transfers a marking m to a marking m', then m' is uniquely determined by the vector equation

$$\mathbf{m}' = \mathbf{m} + \mathbf{t} \ .$$

The sequential behavior of a Petri net with initial marking m_0 is given by its *occurrence sequences.* An occurrence sequence is a finite or infinite sequence of transitions such that m_0 enables the first transition, the reached marking enables the second transition, and so on. The marking m reached by a finite occurrence sequence can be computed using the respective numbers of occurrences of transitions in the sequence: If k_t stands for the number of occurrences of a transition t in the occurrence sequence then the marking m can be computed by

$$\mathbf{m} = \mathbf{m}_0 + \sum_{t \in T} k_t \, \mathbf{t} \ .$$

Usually, this equation is given in its matrix form. The matrix with columns $\mathbf{t}_1 \ldots \mathbf{t}_l$ is called *incidence matrix* of the net N. We will denote the incidence matrix of N by $\mathbf{N}$, i.e., $\mathbf{N} = [\mathbf{t}_1 \cdots \mathbf{t}_l]$. Let $\mathbf{k}$ be the column vector with components $k_{t_1}, \ldots, k_{t_l}$. Then, the matrix form of the above equation reads

$$\mathbf{m} = \mathbf{m}_0 + \mathbf{N} \cdot \mathbf{k} \ .$$

This formula for the marking m is the basis of most concepts mentioned in this contribution. It is a compact way to express the interrelation between markings and numbers of transition occurrences in occurrence sequences. Its linear algebraic background turns out particularly appropriate because it allows to transfer and interpret concepts and results from linear algebra to Petri nets.

We provide one example already at this place: A slight reformulation of the equation given above yields a necessary condition for the a marking m to be reachable from the initial marking m_0 in a net N: the *marking equation* of the marking m,

$$\mathbf{N} \cdot \mathbf{x} = \mathbf{m} - \mathbf{m}_0 \ ,$$

must have a vector solution for $\mathbf{x}$ with components in $\mathbb{N}$. Hence, in particular, there must exist a solution with rational components. A suitable solution vector $\mathbf{k}$ for the variable $\mathbf{x}$ proves that the marking equation is soluble. Conversely, the non-reachability of a marking m can often be shown by proving that the marking equation is not soluble. There are efficient ways to prove non-solubility in the rationals, e.g. Gauss-elimination, but these techniques do not provide short and understandable proofs showing that no solution exists.

Another well-known necessary condition for the reachability of a marking is provided by *place invariants*. A place invariant can be viewed as an integral vector $\mathbf{i}$ satisfying $\mathbf{i} \cdot \mathbf{t} = 0$ for each transition t. Hence, a place invariant vector is a solution $\mathbf{i}$ for $\mathbf{y}$ to the equation system

$$\mathbf{y} \cdot \mathbf{N} = (0, \ldots, 0) .$$

The vectors

$$(1,1,0,0,0,0) \text{ and } (0,0,1,0,1,1)$$

are two place invariant vectors of our example net of Figure 1. It is not difficult to see that, for each marking m reachable from m_0, the products $\mathbf{i} \cdot \mathbf{m}_0$ and $\mathbf{i} \cdot \mathbf{m}$ coincide for any place invariant vector $\mathbf{i}$. Thus, place invariants yield very elegant proofs for the non-reachability of a marking m: every place invariant vector $\mathbf{i}$ satisfying $\mathbf{i} \cdot \mathbf{m}_0 \neq \mathbf{i} \cdot \mathbf{m}$ does the job.

A central theorem of linear algebra is the alternatives' theorem of Fredholm [Schr86]. It implies that the marking equation of a marking m possesses a rational solution if and only if $\mathbf{i} \cdot \mathbf{m}_0 = \mathbf{i} \cdot \mathbf{m}$ for each place invariant vector $\mathbf{i}$. Hence, solubility of the marking equation is shown by a suitable solution vector whereas its non-solubility can be proven by a suitable place invariant vector $\mathbf{i}$ satisfying $\mathbf{i} \cdot \mathbf{m} \neq \mathbf{i} \cdot \mathbf{m}_0$.

Decisive application oriented properties of Petri nets are mostly hard to decide, i.e., there exist no efficient algorithm to decide whether or not the property holds. However, there are necessary or sufficient conditions for many properties, that can efficiently be tested. We distinguish three kinds of techniques in this context:

- a verification technique is a deterministic algorithm to prove a desired property,
- a proof technique is intended to formulate a short proof of a property (formally, it is a non-deterministic algorithm that guesses the proof arguments and then verifies the proof),
- an analysis technique provides a variety of properties of a given marked net, as a basis for further arguments.

For the above discussed property – non-reachability of some marking m – a verification technique finds out that the marking equation has no solution. A proof technique for the same property employs a place invariant. The result of an analysis technique could be a collection of place invariants.

In section 2 we study basic concepts such as equation systems, systems of inequalities, Petri nets and their matrix representations, place and transition invariant vectors. Section 3 concentrates on the marking equation and related necessary conditions for the reachability of a marking. Linear algebraic characterizations of traps and siphons as well as consequences for reachability analysis are the topics of section 4. In section 5, we study state based properties of marked nets. In particular, we identify properties of markings of a net that are satisfied by all reachable markings. We provide a verification technique for such properties, based on the results of the previous sections. Finally, the 6th section is concerned with the so-called rank conditions. These conditions provide necessary criteria and sufficient criteria for liveness of marked nets.

Each section contains bibliographic notes including sources and related concepts.

Bibliographic Notes

The books [Pete81, Reis85, DeEs95] as well as the survey articles [Pete77, JaVa80, Mura89] contain sections on linear algebraic techniques for Petri nets. The papers [MeRo80, Laut87a] are surveys devoted to this topic. References to particular concepts such as place invariant vectors or the marking equation can be found at the end of the respective sections where they are introduced formally.

A more detailed presentation of the results of this contribution and of further applications of linear algebra and linear programming to verification, proofs and analysis of Petri net behavior is provided in the book [Dese98] (in German).

2 Definitions and Elementary Results

2.1 Systems of Linear Inequalities

The set of rational numbers is denoted by $\mathbb{Q}$, the set of integers by $\mathbb{Z}$ and the set of natural numbers (i.e., nonnegative integers) by $\mathbb{N}$. A *positive* integer is greater than 0.

We use bold capitals $\mathbf{A}, \mathbf{B}, \mathbf{C}$ as symbols for matrices, bold $\mathbf{a}, \mathbf{b}, \mathbf{c}$ for vectors and a, b, c for numbers. The transposed of a matrix $\mathbf{A}$ is denoted $\mathbf{A}^\mathsf{T}$, and similarly for vectors. Variable vectors are denoted by $\mathbf{x}$ and $\mathbf{y}$ and variable numbers by x and y. Products of a matrix with some other matrix or vector always assumes proper arities. The product of vectors or matrices with numbers is defined componentwise and is denoted by simple concatenation (without "·").

If $\mathbf{a} = (a_1, \ldots, a_n)$ and $\mathbf{b} = (b_1, \ldots, b_n)$ are vectors, then $\mathbf{a} \le \mathbf{b}$ if $a_i \le b_i$ for all i $(1 \le i \le n)$ and $\mathbf{a} < \mathbf{b}$ if $a_i < b_i$ for all i $(1 \le i \le n)$. We write $\mathbf{a} \ne \mathbf{b}$ if not $\mathbf{a} = \mathbf{b}$. Notice that $\mathbf{a} \le \mathbf{b}, \mathbf{a} \ne \mathbf{b}$ does not imply $\mathbf{a} < \mathbf{b}$. We write $\mathbf{0}$ for vectors with the entry 0 in each component. Each vector $\mathbf{a} > \mathbf{0}$ is *positive*. Each vector $\mathbf{a} \ge \mathbf{0}$ is *nonnegative*. For a set A and a nonnegative integer k, A^k denotes the set of all vectors with k components and all components in A. We write A^* for the set of all vectors over the set A, i.e., vectors with all components in A.

A *linear inequality* is given by an integral vector $\mathbf{a}$ and an integer b. It is represented in the form

$$\mathbf{a} \cdot \mathbf{x} \leq b \, .$$

It is *soluble over a set A* if there exists some $\mathbf{k}$ in A^* satisfying $\mathbf{a} \cdot \mathbf{k} \leq b$.

A *system of linear inequalities* is a set of linear inequalities. It is soluble if there exists a vector that solves all inequalities of the set. If it is finite then it has a matrix based representation

$$\mathbf{A} \cdot \mathbf{x} \leq \mathbf{b} \, ,$$

where the vectors $\mathbf{a}$ of the linear inequalities are the rows of the matrix $\mathbf{A}$ and the numbers b are the components of the vector $\mathbf{b}$. Other representations are:

$\mathbf{A} \cdot \mathbf{x} \geq \mathbf{b}$	for $(-1)\,\mathbf{A} \cdot \mathbf{x} \leq (-1)\,\mathbf{b}$,
$\mathbf{A} \cdot \mathbf{x} \leq \mathbf{b}, \mathbf{C} \cdot \mathbf{x} \leq \mathbf{d}$	for the union of the two sets of inequalities,
$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$	for $\mathbf{A} \cdot \mathbf{x} \leq \mathbf{b}$, $\mathbf{A} \cdot \mathbf{x} \geq \mathbf{b}$.

Thus, each *linear equation system* can be interpreted as a particular system of linear inequalities that contains only equations.

We will frequently employ analysis techniques that are based on the solubility of systems of linear inequalities. The complexity of these techniques depend on the domain under consideration. The following well-known results can e.g. be found in [Schr86].

Proposition 1.

(1) Each system of linear inequalities over $\mathbb{Q}$ can be solved in polynomial time (linear programming).

(2) The solubility of systems of linear inequalities over $\mathbb{Z}$ is NP-complete (integer linear programming, variant 1.)

(3) The solubility of linear equation systems over $\mathbb{N}$ is NP-complete (integer linear programming, variant 2.)

2.2 Petri Nets and Linear Algebraic Representations

For a formal definition of Petri nets and their behavior as well as for usual notations we refer to [DeRe98], in this volume. Here, we only consider nets with finite and nonempty sets of places and transitions. All places and transitions of a net N are said to be *elements* of N. The sets of places, transitions and arcs of a net N are denoted by S_N, T_N and F_N, respectively. The net N is denoted by the triplet (S_N, T_N, F_N). The set of all elements of N is frequently also denoted by N, i.e., $N = S_N \cup T_N$.

For the application of linear algebraic techniques, we assume

$$S = \{s_1, \ldots, s_n\} \text{ and } T = \{t_1, \ldots, t_l\} \, .$$

Then, a mapping $k: S \to A$ can be represented by a row or a column vector $\mathbf{k}$ with n components such that the i-th component of $\mathbf{k}$ is $k(s_i)$ $(1 \leq i \leq n)$.

Such vectors are called *place vectors.* Abusing terminology, we will frequently define place vectors via their associated mappings and write $k(s)$ instead of $\mathbf{k}(s)$. *Transition vectors* and related notions are defined correspondingly.

For example, a marking $m: S \to \mathbb{N}$ of N is represented by the place vector

$$\mathbf{m} = \begin{pmatrix} m(s_1) \\ \vdots \\ m(s_n) \end{pmatrix} \in \mathbb{N}^n .$$

If a transition t occurs at a marking m then the follower marking m' is given by

$$\mathbf{m}' = \mathbf{m} + \mathbf{t}$$

where $\mathbf{t}: S \to \mathbb{N}$ is a place vector defined by

$$\mathbf{t}(s) = \begin{cases} -1 & \text{if } s \in {}^\bullet t \setminus t^\bullet, \\ +1 & \text{if } s \in t^\bullet \setminus {}^\bullet t, \\ 0 & \text{otherwise} . \end{cases}$$

For each place s, the vector $\mathbf{e}_s: S \to \{0,1\}$ is defined by $\mathbf{e}_s(s) = 1$ and $\mathbf{e}_s(s') = 0$ for $s \neq s'$. For a transition t, the transition vector $\mathbf{e}_t$ is defined correspondingly.

For a set A of places, $\chi(A)$ denotes the *characteristic place vector*, defined as the sum of all vectors $\mathbf{e}_s$ for places $s \in A$. *Characteristic transition vectors* of sets of transitions are defined correspondingly.

2.3 Incidence Matrix and Marking Equation

For a net N with n places and l transitions, the $n \times l$-matrix $\mathbf{N} = [\mathbf{t}_1 \cdots \mathbf{t}_l]$ is called *incidence matrix* of N. We will always denote the incidence matrix of a net N by the bold letter $\mathbf{N}$.

For a finite sequence σ of transitions of N, the transition vector $\mathbf{p}_\sigma: T \to \mathbb{N}$ denotes the *Parikh vector* of σ, where $\mathbf{p}_\sigma(t)$ is the number of occurrences of t in the sequence σ.

Each finite occurrence sequence terminates in a marking that can be computed by help of the incidence matrix and the Parikh vector of the sequence:

Theorem 2. *Let $m_0 \xrightarrow{\sigma} m$ be a finite occurrence sequence of a net N. Then*

$$\mathbf{m}_0 + \mathbf{N} \cdot \mathbf{p}_\sigma = \mathbf{m}. \qquad \square$$

The following necessary condition for the reachability of a marking is an immediate consequence:

Corollary 3. *Let N be a net. If a marking m of N is reachable from a marking m_0, then there exists a solution for $\mathbf{x}$ over $\mathbb{N}$ to the equation*

$$\mathbf{N} \cdot \mathbf{x} = \mathbf{m} - \mathbf{m}_0 . \qquad \square$$

For a given net N with initial marking m_0, this equation is called *the marking equation for m.*

Each finite sequence σ of transitions is enabled at some marking (if σ has length k, provide k tokens for each place). However, given a fixed initial marking, no reachable marking may enable σ. Every transition vector over $\mathbb{N}$ can be interpreted as a Parikh vector of a sequence of transitions; hence the following lemma:

Lemma 4. *Let* $\mathbf{k}$ *in* $\mathbb{N}^*$ *be a transition vector of a net* N. *Then there exist markings* m, m' *and an occurrence sequence* $m \xrightarrow{\sigma} m'$ *of* N *such that* $\mathbf{k}$ *is the Parikh vector of* σ. □

2.4 Place Invariants and Transition Invariants

Place invariants have already been mentioned in the introduction. A formal definition was provided in [DeRe98]. An equivalent definition based on linear algebraic concepts is given by the characterization of the following proposition.

Proposition 5. *Let* N *be a net. A mapping* $i: S_N \to \mathbb{Z}$ *is a place invariant if and only if its corresponding row vector* $\mathbf{i}$ *is a solution for* $\mathbf{y}$ *to the homogeneous equation system*

$$\mathbf{y} \cdot \mathbf{N} = \mathbf{0} .$$ □

The vector representation of a place invariant will be called *place invariant vector.*

The fundamental property of place invariants can be rephrased in terms of place invariant vectors. It follows immediately from the marking equation:

Theorem 6. *Let* N *be a net with a place invariant* i. *If a marking* m *is reachable from a marking* m_0, *then*

$$\mathbf{i} \cdot \mathbf{m} = \mathbf{i} \cdot \mathbf{m}_0 .$$

Proof. If m is reachable from m_0, the marking equation for m has some solution. Multiplication of both sides of the equation by $\mathbf{i}$ yields the result. □

Transition invariants and place invariants are defined correspondingly in [DeRe98]. The following proposition provides a linear algebraic characterization of transition invariants:

Proposition 7. *Let* N *be a net. A mapping* $j: T_N \to \mathbb{Z}$ *is a transition invariant if and only if its corresponding column vector* $\mathbf{j}$ *is a solution for* $\mathbf{x}$ *to*

$$\mathbf{N} \cdot \mathbf{x} = \mathbf{0} .$$ □

These transition vectors are called *transition invariant vectors.*

The fundamental property of transition invariants follows immediately from the marking equation:

Theorem 8. *Let N be a net with an occurrence sequence $m \xrightarrow{\sigma} m'$. The Parikh vector $\mathbf{p}_\sigma$ of σ is a transition invariant vector if and only if $m = m'$.*

Proof. $\mathbf{p}_\sigma$ is a transition invariant vector if and only if $\mathbf{N} \cdot \mathbf{p}_\sigma = \mathbf{0}$. This implies the result, because $\mathbf{m}_0 + \mathbf{N} \cdot \mathbf{p}_\sigma = \mathbf{m}$. □

A connectedness theorem was given in [DeRe98]: Each connected net with a live and bounded marking is strongly connected. We finish this section with a similar result:

Theorem 9. *Each connected net with a positive place invariant and a positive transition invariant is strongly connected.*

Proof. Let N be a connected net. Let i be a place invariant of N such that $\mathbf{i} > \mathbf{0}$ and let j be a transition invariant of N such that $\mathbf{j} > \mathbf{0}$. We only prove that, for each arc (u, v) of N, there exists a path from v to u. The results follows by the definition of weak and strong connectedness.

Case 1: $u \in S_N$ and $v \in T_N$.
Let the mapping $j': T \to \mathbb{N}$ be given by

$$j'(t) = \begin{cases} j(t) & \text{if there is a path from } v \text{ to } t, \\ 0 & \text{otherwise.} \end{cases}$$

Consider the place vector $\mathbf{N} \cdot \mathbf{j}'$. Its component for a place s is given by

$$\sum_{u \in {}^\bullet s} j'(u) - \sum_{u \in s^\bullet} j'(u) .$$

We show $\mathbf{N} \cdot \mathbf{j}'(s) \leq 0$ for each place s. Let s be a place.

Assume first $j'(t) = 0$ for each transition t in ${}^\bullet s$. Since $j'(t) \geq 0$ for each transition t by definition,

$$0 = \sum_{t \in {}^\bullet s} j'(t) \leq \sum_{t \in s^\bullet} j'(t) .$$

If $j'(t) = j(t) > 0$ for some transition t in ${}^\bullet s$ then there is a path from v to s. In this case there exists a path from v to any t in $s^\bullet$. Therefore, $j'(t) = j(t) > 0$ for each transition t in $s^\bullet$. So, in this case,

$$0 < \sum_{t \in {}^\bullet s} j'(t) \leq \sum_{t \in {}^\bullet s} j(t) = \sum_{t \in s^\bullet} j(t) = \sum_{t \in s^\bullet} j'(t) .$$

We have shown for both cases:

$$\sum_{t \in {}^\bullet s} j'(u) \leq \sum_{t \in s^\bullet} j'(t) .$$

This finishes the proof of $\mathbf{N} \cdot \mathbf{j}' \leq \mathbf{0}$.

We have $\mathbf{i} \cdot \mathbf{N} = \mathbf{0}$, because i is a place invariant. Therefore, $\mathbf{i} \cdot \mathbf{N} \cdot \mathbf{j}' = 0$. Since $\mathbf{i} > \mathbf{0}$ and $\mathbf{N} \cdot \mathbf{j}' \leq \mathbf{0}$, the vector $\mathbf{N} \cdot \mathbf{j}'$ has no negative component, which implies $\mathbf{N} \cdot \mathbf{j}' = \mathbf{0}$. Hence, j' is a transition invariant.

So we obtain the following inequalities:

$$\begin{aligned} \sum_{t \in {}^\bullet u} j'(t) &= \sum_{t \in u^\bullet} j'(t) && (j' \text{ is a transition invariant}) \\ &\geq j'(v) && (v \in u^\bullet) \\ &= j(v) && (\text{definition of } j') \\ &> 0 && (j \text{ is a positive transition invariant vector})\,. \end{aligned}$$

Hence, there exists a transition $t \in {}^\bullet u$ satisfying $j'(t) > 0$. By the definition of j', there is a path from v to t. This path can be extended by the place u.

Case 2: $u \in T$ and $v \in S$.
Consider the net $N' = (T_N, S_N, F_N)$, where places and transitions of N are swapped. The incidence matrix $\mathbf{N}'$ of N' is given by $(-1)\,\mathbf{N}^{\mathrm{T}}$. So, by Propositions 5 and 7, i is a positive transition invariant of N' and j is a positive place invariant of N'. The arc (v, u) leads from a place of N' to a transition of N'. As shown above (case 1), the net N' contains a path from transition v to place u. In the net N, this path leads from place v to transition u. □

Bibliographic Remarks

Place invariants were first introduced in [LaSc74, Laut75] in a linear algebraic framework. The paper [Ramc74] considers transition invariants (using different notations). Other relevant early references are [Lien76] and [Mura77], where the latter paper concentrates particularly on the marking equation. [Sifa79] gives a proof of the connectedness theorem (Theorem 9) in a more general setting.

3 The Marking Equation

The problem whether a marking m is reachable from a given initial marking m_0 is known to be decidable, but its enormous complexity prevents effective algorithms (see [DeRe98] for references). In this section, we study the inverse problem: is a given marking m not reachable from m_0? As discussed in the introduction, verification and proof techniques based on linear algebra lead to more efficient algorithms that work for many, though not all, instances. Such techniques for verifying or proving non-reachability of a marking are studied in this section.

It was shown in the introduction that there is a solution of the marking equation over the natural numbers for each reachable marking. Thus, the marking equation provides a necessary condition for reachability. This condition can be employed to verify the non-reachability of the marking. Any algorithm can be used that decides integral solubility of the marking equation: The marking under consideration is not reachable in case there is no solution. Nevertheless, there might exist non-reachable markings of a net which do possess an integral solution to the marking equation. Therefore, this verification technique does not work for all non-reachable markings.

Moreover, unfortunately no efficient algorithm is known that decides integral solubility of linear equation systems (Proposition 1). We will consider weaker criteria for reachability, namely solubility of the marking equation over the rational numbers, over the non-negative rational numbers and over the integers. These conditions will lead to more efficient verification algorithms with, admittedly, less expressive power.

As any verification technique, the verification techniques mentioned above can be viewed as proof techniques, too. However, we can do better. As an example, a place invariant can be used to prove the non-reachability of some given marking. It is easy to check that a given place vector is in fact a place invariant vector and that the initial marking and the given marking disagree w.r.t. this invariant. In the introduction, we have shown that a place invariant can be used to prove non-reachability of a marking if and only if the marking equation has no rational solution. Here, we show that more powerful proof techniques correspond to the solubility of the marking equation over the positive rationals and over the integers.

3.1 Solubility of the Marking Equation over the Natural Numbers

The marking equation provides a necessary criterion for the reachability of a marking m: if m is reachable from the initial marking m_0 of a net N then

$$\mathbf{N} \cdot \mathbf{x} = \mathbf{m} - \mathbf{m}_0$$

has a solution for $\mathbf{x}$ in $\mathbb{N}^*$. Hence the following reachability condition for m:

(M1) *The marking equation for m has a solution for* $\mathbf{x}$ *in* $\mathbb{N}^*$.

This condition cannot be decided efficiently in general because (M1) is an instance of "Integer Linear Programming", hance an NP-complete problem (Proposition 1(3)).

The example given in Figure 2 shows that the marking equation provides no sufficient condition for reachability: The marking

$$\mathbf{m}_1 = (0,1,0,0,1,0,0)^{\mathsf{T}}$$

is not reachable from the initial marking

$$\mathbf{m}_0 = (0,0,1,1,0,0,0)^{\mathsf{T}} .$$

However, there is the following solution to the marking equation:

$$\mathbf{N} \cdot (1,1,0,2,2,0,2)^{\mathsf{T}} = \mathbf{m}_1 - \mathbf{m}_0 .$$

All other markings of this net with solvable marking equation are reachable. So the technique only fails for the marking $\mathbf{m}_1$.

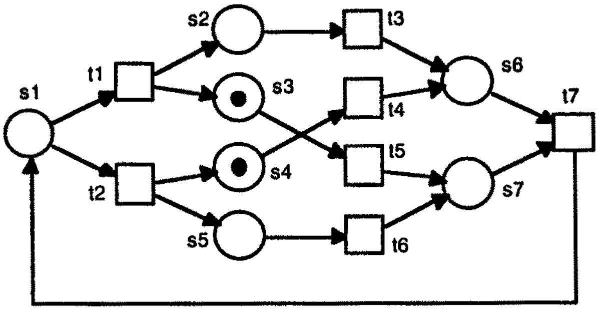


Fig. 2. Limits of the marking equation

The marking equation does not characterize the set of reachable markings because the token count on a place never turns negative. If tokens could be "borrowed" temporarily to avoid negative markings, each marking satisfying the marking equation was reachable. This observation can formally be expressed by addition of another marking at both sides of the marking equation:

Proposition 10. *Let N be a net with markings m_0 and m. If*

$$\mathbf{m}_0 + \mathbf{N} \cdot \mathbf{k} = \mathbf{m}$$

for some $\mathbf{k}$ *in* $\mathbb{N}^*$ *then there exists a marking* $\overline{m}$ *and a transition sequence* σ *such that* $\mathbf{k}$ *is the Parikh vector of* σ *and*

$$\overline{\mathbf{m}} + \mathbf{m}_0 \xrightarrow{\sigma} \overline{\mathbf{m}} + \mathbf{m} \ .$$

Proof. Choose a marking $\overline{m}$ which associates to each place more tokens than the sum of all components of $\mathbf{k}$. □

3.2 Solubility of the Marking Equation over the Rational Numbers

Integral solutions of the marking equation with negative components correspond to "backward occurrences" of transitions. Of course, the marking transformation caused by a transition occurrence is not reversible in general. Another important aspect of the occurrence rule is atomicity of tokens. If tokens could be "broken into pieces", each token could be replaced by suitable "token pieces". If the marking equation has a non-negative rational solution $\mathbf{k}$ with a common denominator d then

$$\mathbf{N} \cdot \mathbf{x}' = d\,\mathbf{m} - d\,\mathbf{m}_0$$

has the non-negative integral solution $d\,\mathbf{k}$ for $\mathbf{x}'$. The vectors $d\,\mathbf{m}_0$ and $d\,\mathbf{m}$ are the initial and the final marking after replacing each token by d "token pieces".

For these reasons, the following condition (M2) properly weakens (M1), because it allows negative and rational numbers.

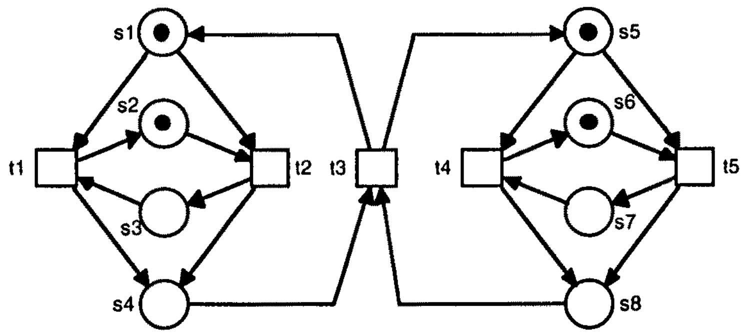


Fig. 3. Limits of rational solutions of the marking equation.

(M2) *The marking equation for m has a solution for* $\mathbf{x}$ *in* $\mathbb{Q}^*$.

The example shown in Figure 3 shows that (M2) has less expressive power than (M1)[1]. The marking

$$\mathbf{m}_1 = (1,0,1,0,1,1,0,0)^\mathsf{T}$$

is not reachable from the depicted initial marking m_0. There is no solution of the marking equation over $\mathbb{Z}$; therefore, there is no solution over $\mathbb{N}$. However, there is a solution over $\mathbb{Q}$, namely

$$\mathbf{m}_0 + \mathbf{N} \cdot \left(0,1,1,\frac{1}{2},\frac{1}{2}\right)^\mathsf{T} = \mathbf{m}_1 .$$

It is easy to see that the marking m_1 would be reachable from m_0 if transitions could occur for "half tokens".

(M2) was discussed in the introduction already. There are efficient algorithms to decide if a given equation system has a rational solution; e.g. the Gauss elimination works in in cubic time. For a related proof technique, there exists a more efficient way using place invariants. The equivalent expressive power of these techniques is stated in the following proposition:

Proposition 11. *Let N be a net with markings* m_0 *and* m. *There exists no solution of the marking equation for* m *over* $\mathbb{Q}$ *if and only if some place invariant* i *satisfies*

$$\mathbf{i} \cdot \mathbf{m}_0 \neq \mathbf{i} \cdot \mathbf{m} .$$

Proof. The proposition follows immediately from Fredholm's theorem of the alternatives (which was first observed by Gauss in the year 1809), see [Schr86, Corollary 3.1b, page 28]. □

[1] This can be shown using simpler examples but we shall need this example later in other contexts, too.

3.3 Solubility of the Marking Equation over the Non-Negative Rational Numbers

Let $\mathbb{Q}_+$ denote the set of non-negative rationals. Then, $\mathbb{Q}_+^*$ contains all rational vectors $\mathbf{x} \geq \mathbf{0}$.

If the marking equation for a marking m has no solution over $\mathbb{Q}_+$, there is in particular no solution over $\mathbb{N}$. So (M1) can be weakened to (M3):

(M3) *The marking equation for* m *has a solution for* $\mathbf{x}$ *in* $\mathbb{Q}_+^*$.

Clearly, this condition is stronger than (M2). It can be expressed by the solubility of the following system of linear inequalities:

$$\begin{aligned} \mathbf{N} \cdot \mathbf{x} &= \mathbf{m} - \mathbf{m}_0 \,, \\ \mathbf{x} &\geq \mathbf{0} \,, \end{aligned}$$

where $\mathbf{x}$ varies over $\mathbb{Q}^*$. This system can be solved in polynomial time, because it is an instance of *Linear Programming* (Proposition 1 (1)). So (M3) provides an efficient verification technique.

For a related *proof technique*, we use the following variant of Farkas' Lemma (see [Schr86, Corollary 7.1c, page 89]):

Proposition 12. *Exactly one of the following systems of inequalities has a solution over* $\mathbb{Q}$*:*

$$\begin{aligned} \mathbf{A} \cdot \mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0} \,, \\ \mathbf{A}^\mathsf{T} \cdot \mathbf{y} \geq \mathbf{0},\ \mathbf{b} \cdot \mathbf{y} < 0 \,. \end{aligned}$$

□

With $\mathbf{N} = \mathbf{A}$ and $\mathbf{b} = \mathbf{m} - \mathbf{m}_0$, (M3) corresponds to the first line of the system given in this proposition. Hence, the complement of (M3) is equivalent to the rational solubility of the following system of inequalities:

$$\begin{aligned} \mathbf{N}^\mathsf{T} \cdot \mathbf{y} &\geq \mathbf{0} \,, \\ (\mathbf{m} - \mathbf{m}_0) \cdot \mathbf{y} &< 0 \,. \end{aligned}$$

Similar to place invariant vectors, any solution of this system of inequalities provides an efficient proof showing that m is not reachable from m_0. Notice that for every rational solution there also is an integral solution, obtained by multiplication with the common denominator.

The example of Figure 3 demonstrates that (M3) has less expressive power than (M1), because the previously given rational solution has no negative components.

If a net has a positive transition invariant, solubility of the marking equation over $\mathbb{Q}$ and solubility of the marking equation over $\mathbb{Q}_+$ coincide, because repeated addition of a positive transition invariant vector to any solution will eventually yield a non-negative solution.

3.4 Solubility of the Marking Equation over the Integers

Another weakening of (M1) is given by

(M4) *The marking equation of the marking m has a solution for* $\mathbf{x}$ *in* $\mathbb{Z}^*$.

The conditions (M4) and (M3) are incomparable. Notice that (M3) and (M4) together do not imply (M1) because the existence of a non-negative solution and the existence of an integral solution does not imply that there is a solution that is both integral and non-negative.

The verification technique based on (M4) corresponds to the integral solubility of the marking equation. The theory of linear diophantic equations implies that the incidence matrix $\mathbf{N}$ can be transformed to its *Smith normal form* (see [Schr86], page 50):

Proposition 13. *Any integral matrix* $\mathbf{N}$ *can be transformed into its Smith normal form* $\mathbf{S}$. *This transformation is given by the matrix product* $\mathbf{S} = \mathbf{Q} \cdot \mathbf{N} \cdot \mathbf{P}$, *where* $\mathbf{Q}$ *and* $\mathbf{P}$ *are unimodular matrices of suitable size.* $\mathbf{S}$ *has the following form:*

$$\mathbf{S} = \begin{pmatrix} s_{1,1} & & & & \\ & \ddots & & 0 & \\ & & s_{a,a} & & \\ & & & 0 & \\ & 0 & & & \ddots \\ & & & & & 0 \end{pmatrix},$$

where $s_{1,1}, \ldots, s_{a,a}$ *are positive integers and, for* $i = 1, \ldots, a-1$, *the diagonal element* $s_{i,i}$ *divides* $s_{i+1,i+1}$. *The Smith normal form is uniquely determined. Its diagonal elements are called elementary divisors of the matrix* $\mathbf{N}$.

$\mathbf{Q}$ *and* $\mathbf{P}$ *are invertible quadratic integral matrices such that their inverses* $\mathbf{Q}^{-1}$ *and* $\mathbf{P}^{-1}$ *are also integral (unimodularity of the matrices).* □

The next theorem shows that integral solubility of the marking equation can be decided by constructing the Smith normal form of the incidence matrix.

Theorem 14. *Let* N *be a net and* m_0, m *be markings of* N. *Let* $\mathbf{S} = \mathbf{Q} \cdot \mathbf{N} \cdot \mathbf{P}$ *be the Smith normal form of* $\mathbf{N}$. *Define*

$$\mathbf{b} = (b_1, b_2, \ldots)^{\mathsf{T}} = \mathbf{Q} \cdot (\mathbf{m} - \mathbf{m}_0) .$$

There exists an integral solution for $\mathbf{x}$ *to the marking equation*

$$\mathbf{N} \cdot \mathbf{x} = \mathbf{m} - \mathbf{m}_0$$

if and only if

(α) $s_{i,i}$ *divides* b_i $(1 \le i \le a)$, *and*
(β) $b_i = 0$ $(i > a)$.

Proof. Since $\mathbf{S} = \mathbf{Q} \cdot \mathbf{N} \cdot \mathbf{P}$, the marking equation can be transformed into

$$\mathbf{Q}^{-1} \cdot \mathbf{S} \cdot \mathbf{P}^{-1} \cdot \mathbf{x} = \mathbf{m} - \mathbf{m}_0 .$$

Define $\mathbf{y} = \mathbf{P}^{-1} \cdot \mathbf{x}$. Multiplication by $\mathbf{Q}$ yields

$$\mathbf{S} \cdot \mathbf{y} = \mathbf{Q} \cdot (\mathbf{m} - \mathbf{m}_0) = \mathbf{b} .$$

Since $\mathbf{P}$ as well as $\mathbf{P}^{-1}$ are integral matrices, this equation system has an integral solution for $\mathbf{y}$ if and only if the marking equation has an integral solution for $\mathbf{x}$. It is easy to see that, by the form of $\mathbf{S}$, there is an integral solution for $\mathbf{y}$ if and only if (α) and (β) hold true. □

The transformation matrix $\mathbf{Q}$ can be constructed in polynomial time [KaBa79]. Given the matrix $\mathbf{Q}$, the conditions (α) and (β) of Theorem 14 can obviously be decided in polynomial time. Hence, by constructing the Smith normal form (including the matrix $\mathbf{Q}$), it can be decided in polynomial time whether the marking equation possesses an integral solution. Thus we obtained an efficient verification technique based on condition (M4).

Place invariants do not suffice for a proof technique based on (M4). They can only prove the non-existence of rational solutions of the marking equation. As mentioned above, in Figure 3, the marking

$$\mathbf{m}_1 = (1, 0, 1, 0, 1, 1, 0, 0)^{\mathsf{T}}$$

agrees with the initial marking on all place invariants but its marking equation has no integral solution.

For this example, it is not difficult to verify that

$$m(\mathtt{s1}) + m(\mathtt{s2}) + m(\mathtt{s5}) + m(\mathtt{s6}) \quad \textit{is an even number}$$

holds for all reachable markings m: If a marking m satisfies the property and the occurrence of a transition leads to m' then m' satisfies the property, too. The same holds for the complementary property (odd token count on the places in $\{\mathtt{s1}, \mathtt{s2}, \mathtt{s5}, \mathtt{s6}\}$). For the markings m_0 and m_1 we obtain the following sums:

$$m_0(\mathtt{s1}) + m_0(\mathtt{s2}) + m_0(\mathtt{s5}) + m_0(\mathtt{s6}) = 4 ,$$

$$m_1(\mathtt{s1}) + m_1(\mathtt{s2}) + m_1(\mathtt{s5}) + m_1(\mathtt{s6}) = 3 .$$

Hence, since this sum is initially even and remains even when transitions occur, m_1 is not reachable from m_0.

The same argument can be stated as follows: For the vector

$$\mathbf{i} = (1, 1, 0, 0, 1, 1, 0, 0)$$

and each column vector $\mathbf{t}$ of the incidence matrix $\mathbf{N}$,

$$\mathbf{i} \cdot \mathbf{t} \equiv 0 \pmod 2 .$$

Each marking transformation $m \xrightarrow{t} m'$ satisfies

$$\mathbf{m}' = \mathbf{m} + \mathbf{t} \ ,$$

where $\mathbf{t}$ is the column of the incidence matrix associated to the transition t. Multiplication by $\mathbf{i}$ yields

$$\mathbf{i} \cdot \mathbf{m}' = \mathbf{i} \cdot \mathbf{m} + \mathbf{i} \cdot \mathbf{t} \ .$$

Since $\mathbf{i} \cdot \mathbf{t} \equiv 0 \pmod 2$,

$$\mathbf{i} \cdot \mathbf{m}' \equiv \mathbf{i} \cdot \mathbf{m} \pmod 2 \ .$$

Vectors with this property will be called modulo place invariant vectors:

Let $k \geq 2$ be a natural number, and let N be a net. A vector $\mathbf{i} \in \mathbb{Z}^*$ is called *modulo-k place invariant vector* of N if, for each column vector $\mathbf{t}$ of $\mathbf{N}$,

$$\mathbf{i} \cdot \mathbf{t} \equiv 0 \pmod k.$$

A vector $\mathbf{i} \in \mathbb{Z}^*$ is called *modulo place invariant vector* if it is a modulo-k place invariant vector for some $k \geq 2$.

Clearly, each place invariant vector is a modulo-k place invariant vector for any k. The inverse does not hold in general. In the above example, the vector $(1,1,0,0,1,1,0,0)$ is a modulo-2 place invariant vector but no (classical) place invariant vector.

Each modulo place invariant vector induces a token conservation law:

Theorem 15. *Let m_0 be an initial marking of a net N and let m be reachable from m_0. Then each modulo-k place invariant vector* $\mathbf{i}$ *of N satisfies*

$$\mathbf{i} \cdot \mathbf{m}_0 \equiv \mathbf{i} \cdot \mathbf{m} \pmod k \ .$$

Proof. Since m is reachable from m_0, there is a solution for $\mathbf{x}$ in $\mathbb{N}^*$ of the marking equation

$$\mathbf{m}_0 + \mathbf{N} \cdot \mathbf{x} = \mathbf{m} \ .$$

Multiplication by $\mathbf{i}$ yields

$$\mathbf{i} \cdot \mathbf{m}_0 + \mathbf{i} \cdot \mathbf{N} \cdot \mathbf{x} = \mathbf{i} \cdot \mathbf{m} \ .$$

An immediate consequence is

$$\mathbf{i} \cdot \mathbf{m}_0 + \mathbf{i} \cdot \mathbf{N} \cdot \mathbf{x} \equiv \mathbf{i} \cdot \mathbf{m} \pmod k \ .$$

Since $\mathbf{i}$ is a modulo-k place invariant vector, all components of $\mathbf{i} \cdot \mathbf{N}$ are multiples of k. Since $\mathbf{x}$ is integral, multiplication by $\mathbf{x}$ yields a multiple of k. Hence,

$$\mathbf{i} \cdot \mathbf{N} \cdot \mathbf{x} \equiv 0 \pmod k \ .$$

So this product can be removed in the above congruence, which completes the proof. □

The inverse of Theorem 15 does not hold true in general. Figure 2 shows a counter example. However, modulo place invariants allow to characterize the integral solubility of the marking equation:

Theorem 16. *Let m_0 and m be markings of a net N. The following propositions are equivalent:*

- *For each integer $k \geq 2$ and each modulo-k place invariant vector $\mathbf{i}$,*
$$\mathbf{i} \cdot \mathbf{m} \equiv \mathbf{i} \cdot \mathbf{m}_0 \pmod{k} .$$
- *The marking equation*
$$\mathbf{m}_0 + \mathbf{N} \cdot \mathbf{x} = \mathbf{m}$$
has a solution for $\mathbf{x}$ over $\mathbb{Z}$.

Proof. The direction ($\Longleftarrow$) was already part of the proof of Theorem 15. So only ($\Longrightarrow$) remains to be proven. Let $\mathbf{S} = \mathbf{Q} \cdot \mathbf{N} \cdot \mathbf{P}$ be the Smith normal form of $\mathbf{N}$ with diagonal elements $s_{1,1}, \ldots, s_{a,a}$. Define $\mathbf{b} = (\mathbf{b_1}, \mathbf{b_2}, \ldots)^{\mathsf{T}} = \mathbf{Q} \cdot (\mathbf{m} - \mathbf{m_0})$. By Theorem 14, it suffices to show:

(α) $s_{i,i}$ divides b_i ($1 \leq i \leq a$), and
(β) $b_i = 0$ ($i > a$).

Let k be the least multiple of $s_{a,a}$ which is greater than all values $|b_i|$, the components of $\mathbf{b}$. By assumption, every modulo-k place invariant vector $\mathbf{i}$ satisfies the congruence
$$\mathbf{i} \cdot (\mathbf{m} - \mathbf{m}_0) \equiv 0 \pmod{k} .$$

(α) Let $1 \leq i \leq a$. Since $s_{i,i}$ divides $s_{a,a}$ and since k is a multiple of $s_{a,a}$, the number $s_{i,i}$ divides k. So we can define the integral vector
$$\mathbf{y}_i = \left(0, \ldots, 0, \frac{k}{s_{i,i}}, 0, \ldots, 0\right)$$
with a positive i-th component. Multiplication by the matrix $\mathbf{S}$ yields
$$\mathbf{y}_i \cdot \mathbf{S} = (0, \ldots, 0, k, 0, \ldots, 0) .$$
Substituting $\mathbf{S}$ by $\mathbf{Q} \cdot \mathbf{N} \cdot \mathbf{P}$ and multiplying both sides of the equation by $\mathbf{P}^{-1}$ proves
$$\mathbf{y}_i \cdot \mathbf{Q} \cdot \mathbf{N} = (0, \ldots, 0, k, 0, \ldots, 0) \cdot \mathbf{P}^{-1} .$$
Since $\mathbf{P}^{-1}$ is an integral matrix, each component on the right hand side of this equation is a multiple of k. Therefore, $\mathbf{y}_i \cdot \mathbf{Q}$ is a modulo-k place invariant vector.

The assumption implies
$$\mathbf{y}_i \cdot \mathbf{Q} \cdot (\mathbf{m} - \mathbf{m}_0) = \mathbf{y}_i \cdot \mathbf{b} \equiv 0 \pmod{k} .$$
By the definition of $\mathbf{y}_i$ follows
$$\frac{k}{s_{i,i}}\, b_i \equiv 0 \pmod{k} .$$

This congruence holds true if and only if $s_{i,i}$ divides b_i.

(β) Let $i > a$. Then

$$\mathbf{e}_i \cdot \mathbf{S} = \mathbf{0} \ .$$

As in the previous case, substituting $\mathbf{S}$ by $\mathbf{Q} \cdot \mathbf{N} \cdot \mathbf{P}$ and multiplying both sides of the equation by $\mathbf{P}^{-1}$ proves

$$\mathbf{e}_i \cdot \mathbf{Q} \cdot \mathbf{N} = \mathbf{0} \ .$$

Hence, $\mathbf{e}_i \cdot \mathbf{Q}$ is a place invariant vector. So it is also a modulo-k place invariant vector. The assumption implies

$$\mathbf{e}_i \cdot \mathbf{Q} \cdot (\mathbf{m} - \mathbf{m}_0) = \mathbf{e}_i \cdot \mathbf{b} = b_i \equiv 0 \ (\mathrm{mod}\ k) \ .$$

Since k was chosen greater than $|b_i|$, we finally obtain $b_i = 0$. □

This theorem states that, if there is no integral solution of the marking equation, then there exists a modulo place invariant proving the non-reachability of the marking under consideration. So the proof technique based on modulo place invariants is exactly as powerful as the verification technique based on (M4).

3.5 Calculating Modulo Place Invariants

Theorem 16 referred to all modulo place invariants of a net. In its proof we showed that it suffices to consider a fixed modulo number k which depends on both the net and the difference between the considered markings. For a given net, no number k suffices for all markings: Every marking $\mathbf{m}_0 + k\,\mathbf{e}_i$ agrees with $\mathbf{m}_0$ on all modulo-k place invariant vectors. However, in general there are no integral solutions for arbitrary vectors $\mathbf{e}_i$ of the corresponding equation $\mathbf{N} \cdot \mathbf{x} = k\,\mathbf{e}_i$.

A finite set of modulo place invariants with different numbers k is not sufficient, too. We could equivalently choose a common multiple of these numbers as a common modulo number and run into the same problem as above. Therefore, no finite set of modulo invariants identifies all markings without integral solution to the marking equation. The situation is different for classical place invariants: Every integral base of the vector space of solutions of $\mathbf{N} \cdot \mathbf{x} = \mathbf{0}$ suffices to prove the non-solubility of the marking equation in the rational numbers.

The following theorem shows that, combining classical place invariants and modulo place invariants, a finite set of vectors comprises the expressive power of all classical place invariants and modulo place invariants.

The following notations will be useful: Each place invariant vector $\mathbf{i}$ of a net N is called a *modulo-0 place invariant vector*. Each place vector $\mathbf{i} \in \mathbb{Z}^*$ is called *modulo-1 place invariant vector*. For arbitrary $u \in \mathbb{Z}$, we write $u \equiv u\ (\mathrm{mod}\ 0)$ and for arbitrary $u, v \in \mathbb{Z}$, we write $u \equiv v\ (\mathrm{mod}\ 1)$.

Proposition 17. *Let u and v be integers, and let $k \in \mathbb{N}$. Then $u \equiv v\ (\mathrm{mod}\ k)$ if and only if $|u - v|$ is a multiple of k.* □

Proposition 18. *Let N be a net, and let $k \in \mathbb{N}$. An integral place vector* $\mathbf{i}$ *is a modulo-k place invariant vector if and only if* $\mathbf{i} \cdot \mathbf{N} = k\,\mathbf{j}$ *for some integral transition vector* $\mathbf{j}$.

Proof. For $k = 0$, the equation coincides with the definition of place invariants. For $k = 1$, choose $\mathbf{j} = \mathbf{i} \cdot \mathbf{N}$. For $k \geq 2$, we have $\mathbf{i} \cdot \mathbf{N} = k\,\mathbf{j}$ for some integral transition vector $\mathbf{j}$ if and only if each component of the vector $\mathbf{i} \cdot \mathbf{N}$ is a multiple of k. Since these components are the products $\mathbf{i} \cdot \mathbf{t}$ for column vectors $\mathbf{t}$ of $\mathbf{N}$, the equation holds exactly for modulo-k place invariant vectors. □

We will show that every row of the transformation matrix $\mathbf{Q}$ of the Smith normal form is a generalized modulo place invariant vector. The so defined set of modulo place invariant vectors is complete in the sense that no larger set allows more implications for the non-reachability of a marking.

Lemma 19. *Let N be a net with n places. Assume that $\mathbf{S} = \mathbf{Q} \cdot \mathbf{N} \cdot \mathbf{P}$ is the Smith normal form of* $\mathbf{N}$ *with diagonal elements $s_{1,1}, \ldots, s_{a,a}$. For $a < i \leq n$, define $s_{i,i} = 0$. Then, for $1 \leq i \leq n$, the i-th row vector of the matrix* $\mathbf{Q}$ *is a modulo-$s_{i,i}$ place invariant vector.*

Proof. For $1 \leq i \leq n$, let $\mathbf{q}_i$ denote the i-th row vector of the matrix Q.

The equation $\mathbf{S} = \mathbf{Q} \cdot \mathbf{N} \cdot \mathbf{P}$ implies

$$\mathbf{Q} \cdot \mathbf{N} = \mathbf{S} \cdot \mathbf{P}^{-1}.$$

Since $\mathbf{S}$ is a diagonal matrix with diagonal entries $s_{1,1}, s_{2,2}, \ldots$, each entry $s_{i,i}$ satisfies

$$\mathbf{q}_i \cdot \mathbf{N} = s_{i,i} \cdot \overline{\mathbf{p}}_i,$$

where $\overline{\mathbf{p}}_i$ is the i-th row of $\mathbf{P}^{-1}$. Since $\mathbf{P}^{-1}$ is an integral matrix, $\overline{\mathbf{p}}_i$ is integral, too. So Proposition 18 can be applied; it yields: $\mathbf{q}_i$ is a modulo-$s_{i,i}$ place invariant vector. □

Theorem 20. *Let m_0 and m be markings of a net N with n places. Assume that $\mathbf{S} = \mathbf{Q} \cdot \mathbf{N} \cdot \mathbf{P}$ is the Smith normal form of* $\mathbf{N}$ *with diagonal elements $s_{1,1}, \ldots, s_{a,a}$. For $a < i \leq n$, define $s_{i,i} = 0$. Then the following propositions are equivalent:*

- *For $1 \leq i \leq n$, the i-th row $\mathbf{q}_i$ of the matrix* $\mathbf{Q}$ *satisfies*

$$\mathbf{q}_i \cdot \mathbf{m} \equiv \mathbf{q}_i \cdot \mathbf{m}_0 \pmod{s_{i,i}}.$$

- *The marking equation*

$$\mathbf{N} \cdot \mathbf{x} = \mathbf{m} - \mathbf{m}_0$$

has an integral solution for $\mathbf{x}$.

Proof.
($\Longleftarrow$) For the rows $\mathbf{q}_i$ satisfying $s_{i,i} = 0$, the vector $\mathbf{q}_i$ is a classical place invariant vector. Since the marking equation has an integral solution, it has in particular a rational solution. So

$$\mathbf{q}_i \cdot \mathbf{m} = \mathbf{q}_i \cdot \mathbf{m}_0 \ .$$

This is equivalent to

$$\mathbf{q}_i \cdot \mathbf{m} \equiv \mathbf{q}_i \cdot \mathbf{m}_0 \ (\text{mod } 0) \ .$$

For the rows $\mathbf{q}_i$ satisfying $s_{i,i} = 1$, nothing has to be shown. For the rows $\mathbf{q}_i$ satisfying $s_{i,i} \geq 2$, the proposition follows immediately from Theorem 16.

($\Longrightarrow$) Let

$$\mathbf{b} = (b_1, \ldots, b_n) = \mathbf{Q} \cdot (\mathbf{m} - \mathbf{m}_0) \ .$$

By Theorem 14, it suffices to show:

(α) $s_{i,i}$ divides b_i ($1 \leq i \leq a$), and
(β) $b_i = 0$ ($i > a$).

The assumption implies that every row $\mathbf{q}_i$ of $\mathbf{Q}$ satisfies

$$\mathbf{q}_i \cdot (\mathbf{m} - \mathbf{m}_0) \equiv 0 \ (\text{mod } s_{i,i}) \ .$$

By the definition of $\mathbf{b}$, b_i is a multiple of $s_{i,i}$ for each i. This implies (α) for $1 \leq i \leq a$, and it implies (β) for $i > a$. □

Bibliographic Remarks

The material of this section is essentially based on [DeNR96], This applies in particular to modulo-invariants and their relation to the marking equation.

A survey on the reachability problem and related topics with emphasis on decidability and complexity issues can be found in [EsNi94]. [Jant87] studies the complexity of solutions of the marking equation in various domains.

[CoSi91] studies non-reachable markings that nevertheless possess solutions to the marking equation.

The reachability problem is simpler for certain subclasses of Petri nets. An example are live and bounded marked free-choice Petri nets where the initial marking can always be reached again. [DeEs93, DeEs95] show that in this class a marking is reachable if and only if the marking equation has a rational solution. So, for this class, place invariants can disprove reachability for all non-reachable markings.

4 Traps and Siphons

The marking equation does not always suffice to identify all non-reachable markings of a marked net. Figure 2 provides an example; the marking

$$\mathbf{m}_1 = (0,1,0,0,1,0,0)^{\mathsf{T}}$$

of the net shown in the figure is not reachable from the depicted initial marking although the marking equation has a nonnegative integral solution.

Another technique for proving non-reachability is given by traps, discussed in detail in [DeRe98]. A trap is a set A of places satisfying $A^{\bullet} \subseteq {}^{\bullet}A$. The following proposition formulates the salient property of traps:

Proposition 21. *Let m_0 and m be markings of a net N such that m is reachable from m_0. Then each trap A of N containing a place marked at m_0 contains some place marked at m.* □

In the net shown in Figure 2, the set

$$A = \{\mathtt{s1}, \mathtt{s3}, \mathtt{s4}, \mathtt{s6}, \mathtt{s7}\}$$

is an initially marked trap. So, for all reachable markings, at least one place of the trap A is marked. In particular, this trap proves that the marking $\mathbf{m}_1$ defined above is not reachable, because it marks no place of A. Thus, traps provide a proof technique for non-reachability that adds expressive power to the techniques based on the marking equation of the previous section.

4.1 Traps and Reachability

Traps generate a necessary condition for a marking m to be reachable from an initial marking m_0 of a net:

> **(T1)** *If some trap A contains a place marked at m_0 then it contains a place marked at m.*

The proof technique for non-reachability based on (T1) formulates a suitable set of places, checks the trap property, and verifies that the initial marking marks at least one place of this trap whereas the marking under consideration does not. Both the defining property of a trap and the conditions for the two markings are easily verified.

It is more difficult to use (T1) for verification. We present a linear algebraic approach in the sequel. A transition might have more input places that belong to some trap than output places. Thus, the token count of a trap can decrease by the occurrence of transitions. The trap condition only requires that, whenever the transition consumes at least one token from a place of a trap then it also adds at least one token to a place of this trap. Hence, by an appropriate weight of output places, we obtain the inequality expressed in the following lemma.

Lemma 22. *A set A of places is a trap if and only if, for each transition t,*

$$|{}^\bullet t| \cdot |A \cap t^\bullet| \geq |A \cap {}^\bullet t| .$$

Proof.
($\Longrightarrow$) Let t be a transition. If $t \notin A^\bullet$ then $|A \cap {}^\bullet t| = 0$ and we are finished. So assume $t \in A^\bullet$. Then $t \in {}^\bullet A$, because A is a trap. Therefore, $|A \cap t^\bullet| \geq 1$. So we obtain

$$|{}^\bullet t| \cdot |A \cap t^\bullet| \geq |{}^\bullet t| \geq |A \cap {}^\bullet t| .$$

($\Longleftarrow$) Let t be a transition in $A^\bullet$. Then $|A \cap {}^\bullet t| > 0$. The hypothesis yields

$$|{}^\bullet t| \cdot |A \cap t^\bullet| \geq |A \cap {}^\bullet t| > 0$$

and therefore $|A \cap t^\bullet| > 0$. This implies $t \in {}^\bullet A$. □

Recall that, for a transition t of a net, $\chi({}^\bullet t)$ is the characteristic vector of its pre-set and $\chi(t^\bullet)$ is the characteristic vector of its post-set. Both vectors are place vectors. Their sum is the column $\mathbf{t}$ of the incidence matrix.

Lemma 23. *A set A of places of a net is a trap if and only if there exists a nonnegative place vector $\mathbf{x}$ in $\mathbb{Q}^*$ such that*

- *a component of $\mathbf{x}$ is 0 if and only if the corresponding place belongs to A (i.e., A is the carrier set of $\mathbf{x}$),*
- *for each transition t of N,*

$$|{}^\bullet t|\, \chi(t^\bullet) \cdot \mathbf{x} \geq \chi({}^\bullet t) \cdot \mathbf{x} .$$

Proof.
($\Longrightarrow$) The characteristic vector of A enjoys the property required for $\mathbf{x}$. We have

$$\chi(t^\bullet) \cdot \chi(A) = |A \cap t^\bullet| \quad \text{and} \quad \chi({}^\bullet t) \cdot \chi(A) = |A \cap {}^\bullet t| .$$

The inequality follows by Lemma 22.

($\Longleftarrow$) Let $t \in A^\bullet$ be a transition. The assumption on $\mathbf{x}$ implies $\chi({}^\bullet t) \cdot \mathbf{x} > 0$. Hence,

$$|{}^\bullet t|\, \chi(t^\bullet) \cdot \mathbf{x} \geq \chi({}^\bullet t) \cdot \mathbf{x} > 0$$

by the second assumption. Therefore, $\chi(t^\bullet) \cdot \mathbf{x} > 0$. Again using the assumption on $\mathbf{x}$, this inequality implies $t \in {}^\bullet A$. Hence A is a trap. □

Deciding (T1) can be reduced by Lemma 23 to the rational solubility of the following system of inequalities:

Theorem 24. *Let N be a net and m_0, m be markings of N. There is a trap*

- *containing a place s satisfying $m_0(s) \geq 1$, and*
- *containing no place s satisfying $m(s) \geq 1$*

if and only if the following system of inequalities has a solution:

$$\begin{aligned} \mathbf{m}_0^{\mathsf{T}} \cdot \mathbf{x} &> 0\,, \\ \mathbf{m}^{\mathsf{T}} \cdot \mathbf{x} &\leq 0\,, \\ (|{}^{\bullet}t| \cdot \chi(t^{\bullet}) - \chi({}^{\bullet}t)) \cdot \mathbf{x} &\geq 0 \quad \textit{for each transition } t\,, \\ \mathbf{x} &\geq \mathbf{0}\,. \end{aligned}$$

Proof. The third and fourth inequality ensure that the carrier set of the solution vector is a trap, by Lemma 23. The first inequality states that initially at least one place of the trap is marked. The second inequality states that no place of the trap is marked at m. □

If the system of inequalities has any solution, then, by Lemma 22, the carrier set of this vector is a trap and the characteristic vector of this trap is a solution, too. So, if soluble at all, the system of inequalities has a solution $\mathbf{x}$ in $\{0,1\}^*$ which moreover satisfies $\mathbf{m}^{\mathsf{T}} \cdot \mathbf{x} = 0$.

(T1) can be decided efficiently, because the system of inequalities of Theorem 24 can be solved in polynomial time, by Proposition 1 (1).

4.2 Siphons and Reachability

Siphons and traps are very similar concepts. Whereas a trap A satisfies $A^{\bullet} \subseteq {}^{\bullet}A$, a siphon is a set of places satisfying ${}^{\bullet}A \subseteq A^{\bullet}$. Siphons containing no initially marked place will never gain a token (see [DeRe98]):

Proposition 25. *Let m_0 and m be markings of a net N such that m is reachable from m_0. Then each siphon A of N containing no place marked at m_0 contains no place marked at m.* □

So also siphons generate a necessary condition for a marking m to be reachable from the initial marking m_0:

(S1) *If some siphon A contains no place marked at m_0 then it contains no place marked at m.*

As for traps, a proof technique for non-reachability based on (S1) formulates a suitable set of places, checks the siphon property, and verifies that the initial marking marks no place of this siphon whereas the marking under consideration marks at least one place of this siphon.

Using (S1) for verification of non-reachability follows closely the above lines for traps. The following lemma is an analogue to Lemma 23:

Lemma 26. *A set A of places of a net is a siphon if and only if there exists a nonnegative place vector* $\mathbf{x}$ *in* $\mathbb{Q}^*$ *such that*

- *a component of* $\mathbf{x}$ *is* 0 *if and only if the corresponding place belongs to* A,
- *for each transition* t *of* N,

$$|t^\bullet|\,\chi({}^\bullet t)\cdot\mathbf{x} \geq \chi(t^\bullet)\cdot\mathbf{x}\,.$$

Proof. By definition, a siphon of a net (S,T,F) is a trap of the net (S,T,F^{-1}) and vice versa, where $F^{-1} = \{(u,v)|(v,u) \in F\}$. The pre-set ${}^\bullet t$ of a transition t in (S,T,F) is equal to the post-set $t^\bullet$ of t in (S,T,F^{-1}) and vice versa. The result follows by Lemma 23. □

By Lemma 26, (S1) reduces to the rational solubility of the following system of inequalities:

Theorem 27. *Let* N *be a net and* m_0, m *be markings of* N. *There is a siphon*

- *containing no place* s *satisfying* $m_0(s) \geq 1$, *and*
- *containing some place* s *satisfying* $m(s) \geq 1$

if and only if the following system of inequalities has a solution:

$$\begin{aligned} \mathbf{m}_0\cdot\mathbf{x} &\leq 0\,,\\ \mathbf{m}\cdot\mathbf{x} &> 0\,,\\ (|t^\bullet|\cdot\chi({}^\bullet t) - \chi(t^\bullet))\cdot\mathbf{x} &\geq 0 \quad \textit{for each transition } t\,,\\ \mathbf{x} &\geq \mathbf{0}\,. \end{aligned}$$

Proof. Analogously to the proof of Theorem 24. □

There is a polynomial time algorithm to decide (S1) by Theorem 27 and Proposition 1 (1).

Bibliographic Remarks

Calculating traps and siphons using equation systems was suggested in [Laut87b]. The characterizations of traps and siphons used in this section are based on results from [EzCS93].

5 Verification of Facts

In contrast to properties like *liveness* and *boundedness*, considered in [DeRe98], properties of Petri nets that are specific for a given net are studied in this section. We concentrate on state based properties, i.e., properties formulated in terms of markings of a net. Figure 4 shows a typical example. This net models an algorithm that guarantees mutual exclusion of critical sections of two concurrent processes. The critical sections are modeled by the places **s2** and **s4**.
For proving mutual exclusion, it has to be shown:

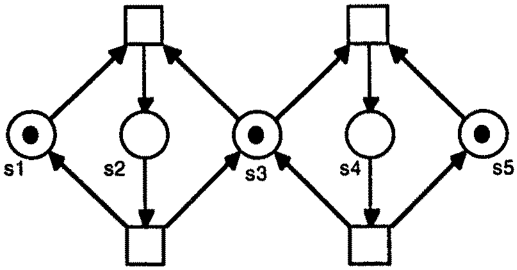


Fig. 4. Mutual exclusion

(M1) *At no reachable marking, both* **s2** *and* **s4** *are marked.*

Conversely, this specification states a property for all reachable markings, namely that either **s2** or **s4** is not marked. It is not difficult to verify that the example net satisfies this specification, by constructing the set of all reachable markings. However, we aim at proving such properties by linear algebraic techniques, because the explicit construction of all reachable markings is not feasible in general.

We restrict to state based properties shaped:

Every reachable marking satisfies φ,

where φ is a predicate over minimal or maximal token counts on places. The property (M1) belongs to this class, because it is equivalently formulated as

(M1′) *Every reachable marking* m *satisfies* $(m(s_2) = 0 \vee m(s_4) = 0)$.

5.1 Linear Predicates

A state of a system corresponds to a marking of its Petri net model. So a *predicate* of a net is a relation on the token count of places. Here, we consider any marking of a net and do not stick to markings reachable from some initial marking. A predicate is interpreted on markings; for each marking, it is either satisfied or not. Hence, a predicate is uniquely determined by the set of markings for which it holds true.

Predicates will be formulated by linear inequalities with variables varying over markings. Every such inequality denotes the predicate given by the set of markings satisfying the inequality. More formally, a predicate φ of a net N is called *linear* if it is satisfied by (and only by) the nonnegative integral solutions for the variable $\mathbf{m}$ of a linear inequality

$$\mathbf{u} \cdot \mathbf{m} \leq v ,$$

where $\mathbf{u}$ is an integral place vector and v is an integer. The predicate φ is said to be *denoted by* this inequality.

We will additionally employ the following notations:

– $\mathbf{u} \cdot \mathbf{m} \geq v$ stands for $(-1)\,\mathbf{u} \cdot \mathbf{m} \leq (-1)\,v$, and
– $\mathbf{u} \cdot \mathbf{m} = v$ represents the two linear predicates $\mathbf{u} \cdot \mathbf{m} \leq v$ and $\mathbf{u} \cdot \mathbf{m} \geq v$.

Notice that variables in these inequalities are markings whereas in previous sections we considered inequalities with transition vectors as variables.

Many important predicates are linear:

(1) If a place s of a net models a logical condition, the corresponding predicate is given by "$m(s) \geq 1$". A corresponding inequality in the above form is

$$\mathbf{e}_s \cdot \mathbf{m} \geq 1 \; .$$

(2) The set of markings that mark at least one place of a set A of places is denoted by the predicate

$$\chi(A) \cdot \mathbf{m} \geq 1 \; .$$

(3) Each place invariant vector $\mathbf{i}$ defines two linear predicates that are satisfied for all markings reachable from the initial marking m_0. They are denoted by

$$\mathbf{i} \cdot \mathbf{m} = (\mathbf{i} \cdot \mathbf{m}_0) \; .$$

(4) An upper bound k of a place s is denoted by

$$\mathbf{e}_s \cdot \mathbf{m} \leq k \; .$$

(5) Two places s and r model mutually exclusive sections if they are never marked together. If moreover both places are 1-bounded, this predicate can be denoted by

$$(\mathbf{e}_s + \mathbf{e}_r) \cdot \mathbf{m} \leq 1 \; .$$

Linear predicates are not closed under conjunction or disjunction in general. The complement of a linear predicate is linear, too:

Proposition 28. *If φ is a linear predicate, denoted by $\mathbf{u} \cdot \mathbf{m} \leq v$ then its complement $\neg\varphi$ is linear; it is denoted by $\mathbf{u} \cdot \mathbf{m} \geq v + 1$.* □

5.2 Implication of Linear Predicates

Implications between linear predicates can be characterized by means of systems of inequalities:

Proposition 29. *Assume linear predicates $\varphi_1, \ldots, \varphi_n$, denoted by inequalities*

$$\mathbf{u}_i \cdot \mathbf{m} \leq v_i \; (1 \leq i \leq n) \; .$$

Let φ be the linear predicate denoted by $\mathbf{u} \cdot \mathbf{m} \leq v$. The set $\{\varphi_1, \ldots, \varphi_n\}$ implies φ (i.e., every marking satisfying all predicates of this set also satisfies φ) if and only if the following inequality system has no solution for $\mathbf{m}$ over $\mathbb{N}$.

$$\begin{aligned} \mathbf{u}_1 \cdot \mathbf{m} &\leq v_1 \; , \\ \mathbf{u}_2 \cdot \mathbf{m} &\leq v_2 \; , \\ &\vdots \\ \mathbf{u}_n \cdot \mathbf{m} &\leq v_n \; , \\ \mathbf{u} \cdot \mathbf{m} &\geq v + 1 \; . \end{aligned}$$

□

In the marked net shown in Figure 2, all reachable markings satisfy the following inequalities:

$$\begin{array}{l} (\,-1, 0, -1, -1, 0, -1, -1\,) \cdot \mathbf{m} \leq -1\,, \\ (\;\;\, 2, 1, \;\;\, 1, \;\;\, 1, 1, \;\;\, 1, \;\;\, 1\,) \cdot \mathbf{m} \leq \;\;\, 2\,. \end{array}$$

Now we show that these predicates imply the mutual exclusion of s_2 and s_5. According to Proposition 29, we add the complement of the inequality

$$(\,0, 1, 0, 0, 1, 0, 0\,) \cdot \mathbf{m} \leq 1$$

to the above inequalities:

$$\begin{array}{l} (\,-1, \;\;\, 0, -1, -1, \;\;\, 0, -1, -1\,) \cdot \mathbf{m} \leq -1, \\ (\;\;\, 2, \;\;\, 1, \;\;\, 1, \;\;\, 1, \;\;\, 1, \;\;\, 1, \;\;\, 1\,) \cdot \mathbf{m} \leq \;\;\, 2, \\ (\;\;\, 0, -1, \;\;\, 0, \;\;\, 0, -1, \;\;\, 0, \;\;\, 0\,) \cdot \mathbf{m} \leq -2. \end{array}$$

All inequalities are formulated in terms of the relation "$\leq$". Hence, summing up the respective left hand sides and right hand sides yields one more inequality which is implied by the other inequalities. In the example, this sum is

$$(\,1, 0, 0, 0, 0, 0, 0\,) \cdot \mathbf{m} \leq -1\,.$$

Clearly, this inequality does not have any nonnegative solution for $\mathbf{m}$. Therefore, the three inequalities above have no common nonnegative solution for $\mathbf{m}$, and we are finished.

The same result is shown easier if the property of the marking m being nonnegative is added in terms of inequalities. Then, it suffices to prove that there is no integral solution for $\mathbf{m}$. In the above example, the inequality

$$(\,-1, 0, 0, 0, 0, 0, 0\,) \cdot \mathbf{m} \leq 0$$

can be added, and we gain the contradiction "$0 \leq -1$".

Proposition 29 also states that a linear predicate φ does not follow from other linear predicates if the inequality system given in the proposition has an integral solution. In fact, every solution constitutes a marking that satisfies all premises but does not satisfy φ.

We mentioned already that neither the conjunction nor the disjunction of linear predicates is linear in general. However, the above proof technique can be extended to such predicates. The following results, recalled from propositional logic, will frequently be used:

Proposition 30. *Let $\varphi_1, \ldots, \varphi_n$ be predicates. The set $\{\varphi_1, \ldots, \varphi_n\}$ implies the conjunction of linear predicates $\psi_1, \ldots, \psi_k$ if and only if it implies every ψ_i $(1 \leq i \leq k)$. It implies the disjunction of linear predicates $\psi_1, \ldots, \psi_k$ if and only if the set $\{\varphi_1, \ldots, \varphi_n, \neg\psi_1, \ldots, \neg\psi_{k-1}\}$ implies the predicate ψ_k.* □

As an example, consider the predicate that holds for all markings of a net except one particular marking m_1 (non-reachability of m_1). This predicate is expressed by a disjunction of linear predicates, denoted by the inequalities

$$\mathbf{e}_s \cdot \mathbf{m} \leq m_1(s) - 1 ,$$
$$\mathbf{e}_s \cdot \mathbf{m} \geq m_1(s) + 1 ,$$

for each place s. For proving that this predicate is implied by a set $\{\varphi_1, \ldots, \varphi_k\}$ of linear predicates, it suffices to show the non-solubility of the following system of inequalities:

$$\begin{aligned} \mathbf{u}_1 \cdot \mathbf{m} &\leq v_1 , \\ &\vdots \\ \mathbf{u}_k \cdot \mathbf{m} &\leq v_k , \\ \mathbf{e}_{s_1} \cdot \mathbf{m} &= m_1(s_1) , \\ &\vdots \\ \mathbf{e}_{s_n} \cdot \mathbf{m} &= m_1(s_n) , \end{aligned}$$

where the first k inequalities denote the predicates $\varphi_1, \ldots, \varphi_k$ and $s_1, \ldots, s_n$ are the places of the net.

Thus, the technique of Proposition 29 can be generalized from linear predicates to propositional formulas constructed from linear predicates. To this end, w.l.o.g. assume a formula in conjunctive normal form. Then, all clauses can be proven as shown above.

5.3 Facts and Place Invariants

We are particularly interested in *facts* of initially marked nets, i.e. predicates that hold for all markings reachable from the initial marking. Facts can be used to establish other facts:

Proposition 31. *Given an initially marked net, each predicate implied by a set of facts is a fact, too.* □

In particular, each weakening of a fact is also a fact.

A given predicate ψ is a fact if and only if the *strongest fact*

$$\psi_0 = \text{"\textit{is reachable from the initial marking}"}$$

implies ψ. In most cases, the proof of this implication is not trivial because ψ_0 characterizes the set of all reachable markings, a set difficult do deal with. Neither the predicate ψ_0 nor the predicate ψ is linear in general. Implications of linear predicates can nevertheless be used, when the following steps are taken.

(1) *The predicate ψ_0 implies linear predicates $\varphi_1, \ldots, \varphi_n$.*
(2) *The set of predicates $\{\varphi_1, \ldots, \varphi_n\}$ implies further linear predicates $\varphi'_1, \ldots, \varphi'_n$.*
(3) *The set of linear predicates $\{\varphi'_1, \ldots, \varphi'_n\}$ implies ψ.*

Step (2) was discussed in the previous subsection. Step (3) consists of logical transformations. As an example, ψ could be the conjunction or the disjunction of linear predicates. In our examples, ψ is always a linear predicate itself. Proof of step (1) requires inequalities from the structure of the net and from the initial marking such that the linear predicates $\varphi_1, \ldots, \varphi_n$ hold true for at least all reachable markings.

Place invariants provide an example, as shown above. In the example of Figure 2, the vector $\mathbf{i} = (2,1,1,1,1,1,1)$ is a place invariant vector. Together with the initial marking, it generates the second of the above inequalities:

$$(2,1,1,1,1,1,1) \cdot \mathbf{m} \leq 2 \ .$$

Other linear predicates are derived from traps; by Proposition 21, every initially marked trap A induces the linear predicate

$$\chi(A) \cdot \mathbf{m} \geq 1 \ .$$

In the above example, the first inequality

$$(-1,0,-1,-1,0,-1,-1) \cdot \mathbf{m} \leq -1$$

states that the set of places $\{s_1, s_3, s_4, s_6, s_7\}$ contains at least one marked place. This predicate is a fact because the set constitutes an initially marked trap. Finally, each initially marked siphon A generates the fact

$$\chi(A) \cdot \mathbf{m} \leq 0 \ .$$

It was previously shown that the marking equation has no solution if and only if some place invariant proves the considered marking non-reachable. Now this result is used for proving facts using place invariants. To this end, let the linear predicates $\varphi_1, \ldots, \varphi_n$ be denoted by the inequalities

$$\mathbf{u}_i \cdot \mathbf{m} \leq v_i \ \ (1 \leq i \leq n) \ .$$

In each inequality, the variable marking $\mathbf{m}$ can be replaced by $\mathbf{m}_0 + \mathbf{N} \cdot \mathbf{x}$ as, for every reachable marking m, there exists some nonnegative integral solution to the marking equation $\mathbf{m}_0 + \mathbf{N} \cdot \mathbf{x} = \mathbf{m}$. Thus, for $1 \leq i \leq n$ we obtain the inequalities

$$\mathbf{u}_i \cdot (\mathbf{m}_0 + \mathbf{N} \cdot \mathbf{x}) \leq v_i \ .$$

Transformation yields

$$(\mathbf{u}_i \cdot \mathbf{N}) \cdot \mathbf{x} \leq v_i - \mathbf{u}_i \cdot \mathbf{m}_0 \ .$$

Each place invariant vector $\mathbf{i}$ induces two linear predicates, given by the equation

$$\mathbf{i} \cdot \mathbf{m} = (\mathbf{i} \cdot \mathbf{m}_0) \ .$$

Substitution into the above inequality leads to

$$(\mathbf{i} \cdot \mathbf{N}) \cdot \mathbf{x} = (\mathbf{i} \cdot \mathbf{m}_0) - \mathbf{i} \cdot \mathbf{m}_0 \ .$$

Since $\mathbf{i}\cdot\mathbf{N} = \mathbf{0}$, both sides evaluate to 0, independently from the value of $\mathbf{x}$. Hence, these inequalities generated by place invariants carry no information and can be dropped. In particular, if a linear predicate follows only from place invariants then it suffices to consider a single inequality:

Theorem 32. *Let N be a net with initial marking m_0. A linear predicate, denoted by $\mathbf{u}\cdot\mathbf{m} \le v$, follows from linear predicates generated by place invariants if and only if the following system of inequalities has no rational solution for $\mathbf{x}$:*

$$(\mathbf{u}\cdot\mathbf{N})\cdot\mathbf{x} \ge v + 1 - \mathbf{u}\cdot\mathbf{m}_0\,,$$
$$\mathbf{x} \ge \mathbf{0}\,.$$

□

5.4 Facts, Traps and Siphons

Each initially marked trap A yields a very simple linear predicate that can be used for proving facts of a marked net:

$$\chi(A)\cdot\mathbf{m} \ge 1,$$

i.e., the trap is marked at each reachable marking m (Proposition 21).

Now we are interested in a corresponding verification technique: Given a linear predicate, is there a trap that proves that this predicate is a fact? In general, the construction and investigation of all traps is not efficient because the number of traps can grow exponentially w.r.t. the size of the net.

The expressive power of all place invariants is characterized by solubility of the marking equation. Hence the verification of a fact using place invariants reduces to the solubility of a single system of inequalities, as shown in the previous theorem. Similarly, we will develop a system of inequalities that has a solution if and only if some trap proves that a given linear predicate is a fact. More formally, we aim at constructing a linear predicate characterizing the set of markings with at least one marked place in each initially marked trap. The inequality for this predicate does not only have variables for markings but also additional variables. The number of these additional variables exceeds the number of transitions of the net by one. Therefore, the solubility of corresponding systems of inequalities is still polynomial in the size of the net.

Recall (T1) from the previous section:

> **(T1)** *If some trap A contains a place marked at the initial marking m_0 then it contains a place marked at m.*

By Lemma 23, a set A is a trap if and only if there exists a nonnegative place vector $\mathbf{x}$ in $\mathbb{Q}^*$ such that

- a component of $\mathbf{x}$ is 0 if and only if the corresponding place belongs to A (i.e., A is the carrier set of $\mathbf{x}$), and
- for each transition t of N we have $|{}^\bullet t|\,\chi(t^\bullet)\cdot\mathbf{x} \ge \chi({}^\bullet t)\cdot\mathbf{x}$.

The vector $|^\bullet t|\,\chi(t^\bullet) - \chi(^\bullet t)$ is a place vector for each transition t. We employ these vectors as columns of a matrix $\mathbf{B}$. Then a set A is a trap if and only if it is the carrier set of a non-negative solution for $\mathbf{x}$ of the inequality

$$\mathbf{B}^{\mathsf{T}} \cdot \mathbf{x} \geq \mathbf{0} .$$

Hence, given a net with an initial marking m_0 and some marking m, there is a trap A containing at least one place marked at m_0 and containing no place marked at m if and only if there is a solution to the following system of inequalities:

$$\begin{aligned} \mathbf{m}_0 \cdot \mathbf{x} &> 0 , \\ \mathbf{m} \cdot \mathbf{x} &\leq 0 , \\ \mathbf{B}^{\mathsf{T}} \cdot \mathbf{x} &\geq \mathbf{0} , \\ \mathbf{x} &\geq \mathbf{0} . \end{aligned}$$

A variant of Farkas' Lemma was employed in the second section of this paper (Proposition 12). Here, another variant will be useful, which can e.g. be found in [Schr86, Corollary 7.1f, page 90].

Proposition 33. *Exactly one of the following systems of inequalities has a solution over* $\mathbb{Q}$*:*

$$\mathbf{A} \cdot \mathbf{y} \leq \mathbf{b},\ \mathbf{y} \geq \mathbf{0} ,$$
$$\mathbf{A}^{\mathsf{T}} \cdot \mathbf{x} \geq \mathbf{0},\ \mathbf{b} \cdot \mathbf{x} < 0,\ \mathbf{x} \geq \mathbf{0} .$$

□

Consider an additional column to the previously defined matrix $\mathbf{B}$ that contains the components of the vector $(-1)\,\mathbf{m}$. Substituting this extended matrix $[\mathbf{B}\,|\,(-1)\,\mathbf{m}]$ for $\mathbf{A}$ and the vector $(-1)\,\mathbf{m}_0$ for $\mathbf{b}$, the above system of inequalities exactly matches the second row of Proposition 33. Therefore, Farkas' Lemma proves that this system has no solution if and only if the following system is soluble:

$$\begin{aligned} [\mathbf{B}\,|\,(-1)\,\mathbf{m}] \cdot \mathbf{y} &\leq (-1)\,\mathbf{m}_0 , \\ \mathbf{y} &\geq \mathbf{0} . \end{aligned}$$

In other words, this system of inequalities has a solution for $\mathbf{y}$ if and only if the marking m marks all initially marked traps.

Let y be the last component of the variable vector $\mathbf{y}$, and decompose $\mathbf{y}$ into $\tilde{\mathbf{y}}\ y$. Then the above system is equivalently represented as:

$$\begin{aligned} y\,\mathbf{m} &\geq \mathbf{B} \cdot \tilde{\mathbf{y}} + \mathbf{m}_0 , \\ \tilde{\mathbf{y}} &\geq \mathbf{0} , \\ y &\geq 0 . \end{aligned}$$

If there is a solution satisfying $y = 0$, the same solution vector $\tilde{\mathbf{y}}$ completed by any positive value for y is also a solution, because $\mathbf{m}$ is a marking with no negative components. So the last line of the above system can be replaced by $y > 0$, without changing the solubility of the system. Division by y yields

$$\begin{aligned} \mathbf{m} &\geq \tfrac{1}{y}\,\mathbf{B} \cdot \tilde{\mathbf{y}} + \tfrac{1}{y}\,\mathbf{m}_0 , \\ \tilde{\mathbf{y}} &\geq \mathbf{0} , \\ y &> 0 . \end{aligned}$$

Setting $\tilde{\mathbf{z}} = \frac{1}{y} \cdot \tilde{\mathbf{y}}$ and $z = \frac{1}{y}$ yields

$$\begin{aligned} \mathbf{m} &\geq \mathbf{B} \cdot \tilde{\mathbf{z}} + z\,\mathbf{m}_0 \ , \\ \tilde{\mathbf{z}} &\geq \mathbf{0} \ , \\ z &> 0 \ . \end{aligned}$$

This form is linear in the variables $\mathbf{m}$, $\tilde{\mathbf{z}}$ and z. The following theorem recaptures the result of these transformations.

Theorem 34. *Let N be a net with initial marking m_0. Each trap containing an initially marked place contains a place marked at a given marking m if and only if the following system of inequalities has a rational solution for $\tilde{\mathbf{z}}$ and z:*

$$[\mathbf{B} \mid \mathbf{m}_0 \mid (-1)\,\mathbf{I}] \cdot \begin{bmatrix} \tilde{\mathbf{z}} \\ z \\ \mathbf{m} \end{bmatrix} \geq 0 \ , \ \tilde{\mathbf{z}} \geq \mathbf{0} \ , \ z > 0 \ ,$$

where the matrix $\mathbf{B}$ is the matrix defined above and $\mathbf{I}$ is the identity matrix. □

Since, for every reachable marking, there is a solution to this system of inequalities, every implied linear predicate is a fact. Its expressive power comprises the expressive power of all initially marked traps:

Corollary 35. *Given an initially marked trap A, its generated fact $\chi(A) \cdot \mathbf{m} > 0$ is implied by the system of inequalities given in Theorem 34.* □

Siphons can be used for the verification of facts in a similar manner. Recall the reachability criterion for a marking m generated by siphons from the previous section:

> **(S1)** *If some siphon A contains no place marked at the initial marking m_0 then it contains no place marked at m.*

By Lemma 26, a set of places A is a trap if and only if there exists a non-negative place vector $\mathbf{x}$ in $\mathbb{Q}^*$ such that

- a component of $\mathbf{x}$ is 0 if and only if the corresponding place belongs to A,
- for each transition t of N, we have $|t^\bullet|\, \chi({}^\bullet t) \cdot \mathbf{x} \geq \chi(t^\bullet) \cdot \mathbf{x}$.

Similarly as above, we employ the vectors $|t^\bullet|\chi({}^\bullet t) - \chi(t^\bullet)$ as columns of a matrix $\mathbf{C}$. Then a set A is a siphon if and only if it is the carrier set of a non-negative solution for $\mathbf{x}$ of the inequality

$$\mathbf{C}^{\mathsf{T}} \cdot \mathbf{x} \geq \mathbf{0} \ .$$

Hence, given a net with an initial marking m_0 and some marking m, there is a siphon A containing at least one place marked at m and containing no place marked at m_0 if and only if there is a solution to the following system of inequalities:

$$\begin{aligned} \mathbf{m} \cdot \mathbf{x} &> 0\,, \\ \mathbf{m}_0 \cdot \mathbf{x} &\leq 0\,, \\ \mathbf{C}^{\mathsf{T}} \cdot \mathbf{x} &\geq \mathbf{0}\,, \\ \mathbf{x} &\geq \mathbf{0}\,. \end{aligned}$$

We apply again Farkas' Lemma (Proposition 33), taking $[\mathbf{C}\,|\,(-1)\,\mathbf{m}_0]$ for $\mathbf{A}$ and $(-1)\,\mathbf{m}$ for $\mathbf{b}$. Thus, there is a solution of the above system if and only if the following system of inequalities has no solution:

$$\begin{aligned} [\mathbf{C}\,|\,(-1)\,\mathbf{m}_0] \cdot \mathbf{y} &\leq (-1)\,\mathbf{m}\,, \\ \mathbf{y} &\geq \mathbf{0}\,. \end{aligned}$$

The following theorem follows immediately:

Theorem 36. *Let N be a net with initial marking m_0. Each trap containing no initially marked place contains no place marked at a given marking m if and only if the following system of inequalities has a rational solution for $\tilde{\mathbf{y}}$ and y:*

$$[\mathbf{C}\,|\,(-1)\,\mathbf{m}_0\,|\,\mathbf{I}] \cdot \begin{bmatrix} \tilde{\mathbf{y}} \\ y \\ \mathbf{m} \end{bmatrix} \leq 0\,,\ \tilde{\mathbf{y}} \geq \mathbf{0}\,,\ y \geq 0\,,$$

where $\mathbf{C}$ *is the matrix defined above and* $\mathbf{I}$ *is the identity matrix.* □

This system of inequalities has a solution for every reachable marking because unmarked siphons remain unmarked. Its expressive power comprises the expressive power of all initially unmarked siphons:

Corollary 37. *Given an initially unmarked siphon A, its generated fact $\chi(A) \cdot \mathbf{m} \leq 0$ is implied by the system of inequalities given in Theorem 36.* □

5.5 Deadlock-freeness and Facts

The enabling condition of a transition can be formulated by means of a conjunction of linear predicates: For each input place of the transition, a linear predicate demands a token. For transitions with only one input place, enabledness is thus described by exactly one linear predicate, which is called *enabling predicate* of the transition.

If a transition has more than one input place but all its input places are bounded by 1, then one enabling predicate is sufficient, too:

Proposition 38. *Let t be a transition of a marked net and assume that every place in ${}^\bullet t$ is 1-bounded. Then t is enabled at a reachable marking m if and only if*

$$\chi({}^\bullet t) \cdot \mathbf{m} \geq |{}^\bullet t|\,.$$

Proof. The value $\chi({}^\bullet t) \cdot \mathbf{m}$ denotes the number of tokens on input places of t at a marking m. If none of these places carry more than one token, $\chi({}^\bullet t) \cdot \mathbf{m} \geq |{}^\bullet t|$ implies that every input place is marked at m. Conversely, $\chi({}^\bullet t) \cdot \mathbf{m} < |{}^\bullet t|$ implies that at least one place in ${}^\bullet t$ is unmarked at m. □

By 1-boundedness, the value $\chi({}^\bullet t) \cdot \mathbf{m}$ never exceeds $|{}^\bullet t|$. Hence, the relation symbol "$\geq$" can be replaced by "$=$".

If every transition of a net has a single enabling predicate, a marking is dead if and only if the disjunction of these enabling predicates holds true for this marking. Hence, for showing deadlock-freeness, this disjunction has to be proven a fact. The following theorem states this result for 1-bounded marked nets:

Theorem 39. *A net N with a 1-bounded initial marking m_0 is deadlock-free if and only if every reachable marking m satisfies the linear predicate*

$$\chi({}^\bullet t) \cdot \mathbf{m} \geq |{}^\bullet t|$$

for at least one transition t. □

In all previous examples, it is possible to prove deadlock-freeness with this technique, employing only place invariants. In other words, in all these marked nets, the disjunction of the enabling predicates is implied by linear predicates generated by place invariants.

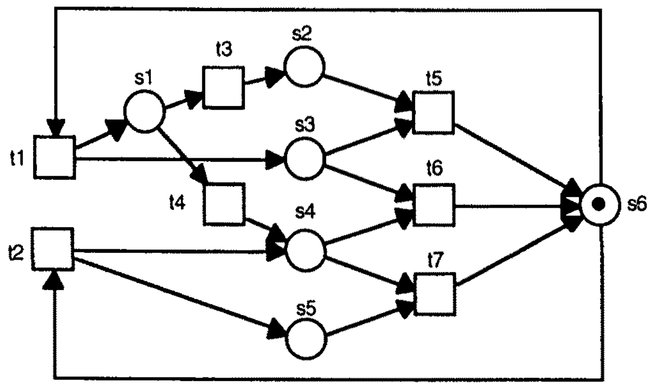


Fig. 5. A deadlock-free marked net

The marked net shown in Figure 5 is also deadlock-free but this property cannot be proven by help of place invariants only. The marking

$$\mathbf{m} = (0, 1, 0, 0, 1, 0)^\mathsf{T}$$

is a (non-reachable) dead marking. However, we have

$$\mathbf{m}_0 + \mathbf{N} \cdot (2,2,2,0,1,1,1)^{\mathsf{T}} = \mathbf{m} \ .$$

Hence, by Proposition 11, the marking m agrees with the initial marking m_0 on all place invariants. However, non-reachability of m can be proven by help of a trap; the set

$$A = \{\mathtt{s3}, \mathtt{s4}, \mathtt{s6}\}$$

constitutes an initially marked trap which is not marked at m.

Not only the dead marking m is non-reachable, but any dead marking, as will be shown next. Consider the place invariant vectors

$$\mathbf{i}_1 = (1,1,0,1,0,1) \text{ and } \mathbf{i}_2 = (0,0,1,0,1,1) \ .$$

The place invariant vector $\mathbf{i}_1$ proves that the places **s1**, **s2**, **s4** and **s6** are 1-bounded, whereas $\mathbf{i}_2$ moreover proves 1-boundedness of the places **s3** and **s5**. We show that these place invariants and the trap A suffice to prove that the disjunction of the enabling predicates is a fact. Equivalently, we show that the following system of inequalities has no integral solution for $\mathbf{m}$:

[1]	$(1,1,0,1,0,1) \cdot \mathbf{m} = 1$	place invariant $\mathbf{i}_1$,
[2]	$(0,0,1,0,1,1) \cdot \mathbf{m} = 1$	place invariant $\mathbf{i}_2$,
[3]	$(0,0,1,1,0,1) \cdot \mathbf{m} \geq 1$	trap A ,
[4]	$(0,0,0,0,0,1) \cdot \mathbf{m} = 0$	**t1** and **t2** are not enabled,
[5]	$(1,0,0,0,0,0) \cdot \mathbf{m} = 0$	**t3** and **t4** are not enabled,
[6]	$(0,1,1,0,0,0) \cdot \mathbf{m} \leq 1$	**t5** is not enabled,
[7]	$(0,0,1,1,0,0) \cdot \mathbf{m} \leq 1$	**t6** is not enabled,
[8]	$(0,0,0,1,1,0) \cdot \mathbf{m} \leq 1$	**t7** is not enabled,

Assume that some marking m satisfies all these inequalities; we aim at deriving a contradiction.

By [5], $m(\mathtt{s1}) = 0$. By [4], $m(\mathtt{s6}) = 0$.

Assume that $m(\mathtt{s3}) \geq 1$. Then, by [6], $m(\mathtt{s2}) = 0$ and, by [7], $m(\mathtt{s4}) = 0$, contradicting [1].

Now assume that $m(\mathtt{s3}) = 0$. Then, by [3], $m(\mathtt{s4}) \geq 0$ and hence, by [8], $m(\mathtt{s5}) = 0$, contradicting [2].

So we obtain a contradiction in both cases, which finishes the proof.

Notice that, for this example, we have to consider solubility of the system of inequalities over $\mathbb{Z}$ because there exists a rational solution; the marking

$$\mathbf{m} = \left(0, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 0\right)^{\mathsf{T}}$$

satisfies all above inequalities.

Bibliographic Remarks

Place invariants and other structural techniques for proving facts are suggested in many publications. See e.g. [Reis85]. The main results on traps (Subsection 5.4) are taken from [EsMe96]. The representation of the enabling condition by a linear inequality was introduced in [Dese85]. The example of a live and safe marked net for which deadlock-freeness can not be proved by place invariants can be found in [Dese88]. The paper [TeCS93] contains a deeper discussion of this analysis technique. In particular, techniques for reducing the number of enabling predicates are presented.

6 The Rank Conditions

The rank conditions provide a sufficient criterion and a necessary criterion for liveness of bounded marked Petri nets. In particular, they establish relations between a net's behavioral property and the rank of its incidence matrix. This connection might be surprising because the rank of the incidence matrix has no obvious interpretation in terms of behavior. Moreover, it is invariant against addition of rows or columns, multiplication by numbers and even transposition. All the corresponding transformations of a net structure cause significant changes of behavior that do not preserve properties such as liveness.

As shown in [DeRe98], every live and bounded marked net has a positive transition invariant. The column rank of the incidence matrix of a live and bounded marked net is therefore not maximal; it never exceeds $|T| - 1$, where T is the set of transitions of the net. Transition invariants are closely related to occurrence sequences that lead from a marking back to itself. If a net has only one transition invariant (and its multiples) then the relative number of transition occurrences in an occurrence sequence that leads from a marking to itself is very restricted. In particular, no conflict can occur again and again. If there is a conflict between transitions t_1 and t_2 then t_1 could be chosen more often than t_2 or vice versa. In this case, there exist transition invariant vectors $\mathbf{j}_1 \geq \mathbf{0}$ and $\mathbf{j}_2 \geq \mathbf{0}$ such that

$$\mathbf{j}_1(t_1) > \mathbf{j}_1(t_2) \quad \text{and} \quad \mathbf{j}_2(t_1) < \mathbf{j}_2(t_2).$$

Then the rank of the incidence matrix does not exceed $|T| - 2$ because $\mathbf{j}_1$ and $\mathbf{j}_2$ are linearly independent.

These considerations suggest that the rank of the incidence matrix of a net is related to the number of its possible conflict situations. The rank conditions formulate such relations.

Two distinct transitions are said to be in *potential conflict* if their pre-sets are not disjoint. The conflict appears when both transitions are enabled and a common input place carries only one token. The *conflict area* of a transition t is the minimal set of transitions that contains t and is closed under potentially conflicting tansitions. Formulated differently, it is an equivalence class of the

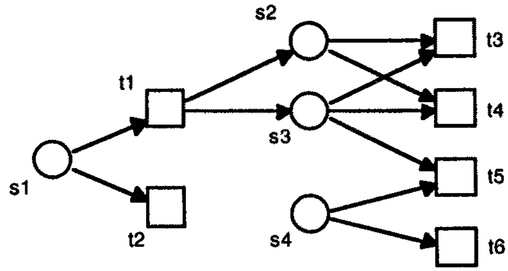


Fig. 6. A net with two conflict areas and four different pre-sets of transitions

transitive closure of potential conflict. In the example of Figure 6, the conflict areas are {t1, t2} and {t3, t4, t5, t6}.

The *sufficient rank condition* requires the rank of the incidence matrix to be smaller than the number of conflict areas of the net.

In the sequel, we consider only nets with each transition having a nonempty pre-set. Two transitions are in *independent conflict* if they have identical pre-sets. Every marking enabling one of the transitions enables the other one, too. This relation is an equivalence relation. The number of equivalence classes equals the number of different pre-sets of transitions. In the example of Figure 6, these different pre-sets are {s1}, {s2, s3}, {s3, s4} and {s4}.

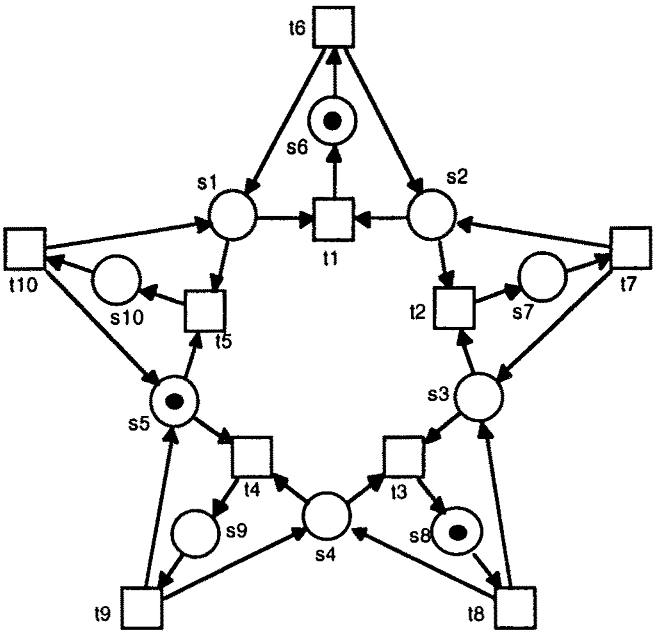
The *necessary rank condition* requires the rank of the incidence matrix to be smaller than the number of different pre-sets of transitions.

In the first subsection, *strong occurrence sequences* will be introduced. This technical notation will be used to prove the sufficient rank condition. Strong occurrence sequences obey a more restrictive enabling rule in the occurrence condition. *Strongly live markings* are based on strong occurrence sequences. The second subsection considers the sufficient rank equation. This condition is sufficient for the liveness of a marking. It is sufficient and necessary for strong liveness of a marking, as will be shown in the third subsection. In the fourth subsection, the necessary rank condition will be proven. We do not only provide the upper bound for the rank of this condition but a lower bound, too.

We stick to nets with positive place and transition invariants. The existence of a positive place invariant implies that every marking is bounded. If a net with a positive place invariant can be marked lively, then it even has a live and bounded marking. In this case, there also exists a positive transition invariant (see [DeRe98]). Moreover, each net with both a positive place invariant and a positive transition invariant is strongly connected (Theorem 9).

We provide a sufficient condition for liveness of a marked net with positive place and transition invariants. First we give a condition for *potential liveness* of a net:

If the rank of the incidence matrix is smaller than the number of conflict areas then the net can be marked lively.



	t1	t2	t3	t4	t5	t6	t7	t8	t9	t10
s1	−1	0	0	0	−1	1	0	0	0	1
s2	−1	−1	0	0	0	1	1	0	0	0
s3	0	−1	−1	0	0	0	1	1	0	0
s4	0	0	−1	−1	0	0	0	1	1	0
s5	0	0	0	−1	−1	0	0	0	1	1
s6	1	0	0	0	0	−1	0	0	0	0
s7	0	1	0	0	0	0	−1	0	0	0
s8	0	0	1	0	0	0	0	−1	0	0
s9	0	0	0	1	0	0	0	0	−1	0
s10	0	0	0	0	1	0	0	0	0	−1

Fig. 7. An example net and its incidence matrix

Before proving this result, we demonstrate it by help of an example. Consider the net shown in Figure 7. It has the positive place invariant vector

$$\mathbf{i} = (1,1,1,1,1,2,2,2,2,2)$$

and the positive transition invariant vector

$$\mathbf{j} = (1,1,1,1,1,1,1,1,1,1).$$

The rank of the incidence matrix is 5. The set of vectors $\{\mathtt{t1},\mathtt{t2},\mathtt{t3},\mathtt{t4},\mathtt{t5}\}$ is a basis of its columns.

The net has 6 conflict areas,

$$\{\mathtt{t1},\mathtt{t2},\mathtt{t3},\mathtt{t4},\mathtt{t5}\},\{\mathtt{t6}\},\{\mathtt{t7}\},\{\mathtt{t8}\},\{\mathtt{t9}\},\{\mathtt{t10}\}.$$

So the premise of the above condition is fulfilled. Hence, the result states that the net can be marked lively. In fact, there are live markings; an example is the marking depicted in the figure.

6.1 Strong Occurrence Sequences and Strongly Live Markings

Two transitions are said to be in a *conflict situation* at a marking if both transitions are enabled and share a common input place. Conversely, a common input place does not guarantee that both transitions are only enabled in conflict situations. For example, if one of the transition becomes enabled before the other transition then it might occur before the other transition is enabled. Such a *confused* situation can sometimes be avoided by the following restricted enabling condition:

> *A transition can occur only in case all transitions of its conflict area are enabled, i.e., all places in the pre-set of transitions of its conflict area are marked.*

In this case, the transition is called *strongly enabled.*

Not every transition enabled at a marking is strongly enabled; a marking might enable only a subset of a conflict area. In the example, transition `t6` is strongly enabled, because {`t6`} is a conflict area. The marking obtained after the occurrence of `t6` marks the places `s1`, `s2`, `s5` and `s8`. It enables, among other transitions, `t1`. However, it does not strongly enable `t1` because `t2` belongs to the same conflict area and is not enabled.

A finite occurrence sequence $m_0 \xrightarrow{\sigma} m_n$ is called a *strong occurrence sequence* if, with $\sigma = t_1 \ldots t_n$ and

$$m_0 \xrightarrow{t_1} m_1 \xrightarrow{t_2} \cdots \xrightarrow{t_n} m_n \quad ,$$

each marking m_{i-1} strongly enables the transition t_i $(1 \le i \le n)$. In the example, the sequence

$$\sigma = \mathtt{t6}\ \mathtt{t8}\ \mathtt{t1}$$

is a strong occurrence sequence, enabled at the initial marking.

A marking m_0 is *strongly live* if, for each reachable marking m and every transition t, there exists a strong occurrence sequence $m \xrightarrow{\sigma} m'$ such that t occurs in σ. The initial marking of the above example net is strongly live.

Clearly, each strongly enabled transition is enabled and each strong occurrence sequence is an occurrence sequence. Each strongly live marking is live, because, in the definition of strong liveness, the marking m is any reachable marking.

6.2 A Sufficient Condition for the Existence of a Live Marking

A net is potentially live if there exists a live initial marking. In this subsection we show that a net with positive place and transition invariants is potentially live if the rank of its incidence matrix is smaller than the number of its conflict areas. Actually, we show the contraposition: if the net has no live initial marking then the rank of its incidence matrix is not smaller than the number of conflict areas. The proof consists of the following steps:

(1) *A given net N with positive place and transition invariants has no live initial marking.*

The definition of strong liveness and (1) imply:

(2) *The net N has no strongly live initial marking.*

The following Lemma 40 will show that (2) implies:

(3) *For each initial marking of N, there is a reachable marking which does not strongly enable any transition.*

This result will be applied to the marking that associates one token to each place. Then the following Lemma 41 will prove:

(4) *There exists a set of transitions $\{t_1, \dots, t_n\}$ of N, containing exactly one transition of each conflict area, and a place vector $\mathbf{x}$ satisfying $\mathbf{x} \cdot \mathbf{t}_i < 0$ $(1 \leq i \leq n)$.*

Consider the matrix $[\mathbf{t}_1 \dots \mathbf{t}_n]$. By (4), the vector $\mathbf{x} \cdot [\mathbf{t}_1 \dots \mathbf{t}_n]$ has only negative components. Assume a solution $\mathbf{j}$ of $[\mathbf{t}_1 \dots \mathbf{t}_n] \cdot \mathbf{y} = 0$. Then either $\mathbf{j} = \mathbf{0}$ or $\mathbf{j}$ has both positive and negative components. In particular, we obtain:

(5) *Every solution $\mathbf{j}$ of $[\mathbf{t}_1 \dots \mathbf{t}_n] \cdot \mathbf{y} = 0$, $\mathbf{y} \geq \mathbf{0}$ satisfies $\mathbf{j} = \mathbf{0}$.*

Together with (5), Lemma 42 will prove:

(6) *The set $\{\mathbf{t}_1, \dots, \mathbf{t}_n\}$ is linearly independent.*

Finally, (6) immediately implies the result we are after:

(7) *The rank of* $\mathbf{N}$ *is not smaller than the number of conflict areas of N.*

Lemma 40. *Let N be a net with positive place and transition invariants and let m_0 be an initial marking of N which is not strongly live. Then there is a reachable marking m such that no transition is strongly enabled at m.*

Proof. Since m_0 is not strongly live, there exist a reachable marking m and a transition t such that m does not enable any strong occurrence sequence containing t. In this situation, t is said to be *excluded* at m. Every transition excluded at a marking m remains excluded at any marking reachable from m by a strong occurrence sequence. Now w.l.o.g. assume that m excludes a maximal set of transitions. Then all markings reachable from m by strong occurrence sequences exclude the same set of transitions.

If a transition is excluded at a marking m then so are all transitions of its conflict area. Let K be a conflict area such that all its transitions are excluded at m. Define $R = {}^\bullet K$. The net N is strongly connected because it has positive place and transition invariants (Theorem 9). Hence, R is not the empty set. The definition of conflict areas implies $K = R^\bullet$. Therefore, no strong occurrence sequence enabled at m contains transitions in $R^\bullet$.

The marking m is bounded because N has a positive place invariant. Hence, the number of tokens on places in R cannot grow unlimited and transitions in ${}^\bullet R$ can occur only finitely often in strong occurrence sequences enabled at m. Consider a strong occurrence sequence which is enabled at m and leads to some marking that also excludes all transitions in ${}^\bullet R$. We have chosen the marking m such that the number of excluded transitions is maximal at m. Hence, all transitions in ${}^\bullet R$ are already excluded at m.

Now assume that u is a transition excluded at m. Then u belongs to some conflict area K. We have shown that every transition in K and every transition in ${}^\bullet({}^\bullet K)$ is excluded at m. So every transition connected to u by a directed path ending at u is excluded at m. There exists at least one transition excluded at m, vic. transition t. Therefore, since N is strongly connected, every transition of N is excluded at m. In other words: no transition is strongly enabled at m. □

Lemma 41. *Let N be a net and let m_0 be the marking of N that associates one token to each place. Assume an occurrence sequence $m_0 \xrightarrow{\sigma} m$ such that no transition of N is strongly enabled at m. Then, for every set $\{K_1, \ldots, K_k\}$ of conflict areas of N there are transitions $t_1 \in K_1, \ldots, t_k \in K_k$ and a place vector $\mathbf{x}$ in $\mathbb{N}^*$ satisfying $\mathbf{x} \cdot \mathbf{t}_i < 0$ $(1 \le i \le k)$.*

Proof. We proceed by induction on k, the number of considered conflict areas.

Base. If $k = 0$, nothing has to be shown.

Step. Assume $k \ge 1$ and define $K = K_1 \cup \cdots \cup K_k$. Let R denote the set of places in ${}^\bullet K$ that are not marked at m. The definition of a conflict area implies $R^\bullet \subseteq K$.

No transition is strongly enabled at the marking m. Hence, every conflict area in $\{K_1, \ldots, K_k\}$ has at least one input place in R. Since m_0 marks all places, the occurrence sequence σ contains transitions in $R^\bullet$ (at least one for each conflict area). Assume that t is the last transition occurring in σ that belongs to $R^\bullet$. Then $t \notin {}^\bullet R$ because places in R are not marked at m.

W.l.o.g. assume that t belongs to the conflict area K_k. We apply the induction hypothesis to $\{K_1, \ldots, K_{k-1}\}$ and obtain transitions $t_1 \in K_1, \ldots, t_{k-1} \in K_{k-1}$ and a place vector $\mathbf{x}'$ in $\mathbb{N}^*$ such that $\mathbf{x}' \cdot \mathbf{t}_i < 0$ $(1 \le i \le k-1)$. Now choose the transition t for the conflict area K_k, i.e., set $t_k := t$. The place vector $\mathbf{x}$ is defined by

$$\mathbf{x}(s) = \begin{cases} (|R|+1)\,\mathbf{x}'(s) & \text{if } s \in {}^\bullet(K_1 \cup \cdots \cup K_{k-1}) \cap R, \\ 1 & \text{if } s \in {}^\bullet K_k \cap R, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly all components of $\mathbf{x}$ are nonnegative integers. Moreover, every positive component belongs to a place in R. We have to show that $\mathbf{x} \cdot \mathbf{t}_i < 0$ $(1 \le i \le k)$.

Case 1. $i = k$. Then $t_i = t_k = t$. Recall that the vector $\mathbf{t}(s)$ equals 1 for $s \in t^\bullet \setminus {}^\bullet t$, -1 for $s \in {}^\bullet t \setminus t^\bullet$ and 0 otherwise. So

$$\mathbf{x} \cdot \mathbf{t} = \sum_{s \in t^\bullet} \mathbf{x}(s) - \sum_{s \in {}^\bullet t} \mathbf{x}(s) \ .$$

We have $t^\bullet \cap R = \emptyset$ because $t \notin {}^\bullet R$. So the first sum equals 0. Since $t \in K_k$ and $t \in R^\bullet$, there exists a place s in ${}^\bullet t$ satisfying $s \in {}^\bullet K_k \cap R$. Then $\mathbf{x}(s) = 1$. So the second sum evaluates to at least 1, and we obtain $\mathbf{x} \cdot \mathbf{t} \le -1$.

Case 2. For $1 \le i \le k-1$, the definition of $\mathbf{x}$ implies:

$$\mathbf{x} \cdot \mathbf{t}_i \ \le \ (|R|+1)\,\mathbf{x}' \cdot \mathbf{t}_i + |t_i^\bullet \cap ({}^\bullet K_k \cap R)| \ \le \ (|R|+1)\,\mathbf{x}' \cdot \mathbf{t}_i + |R| \ .$$

Since $\mathbf{x}' \cdot \mathbf{t}_i$ is negative and integral, we have $\mathbf{x}' \cdot \mathbf{t}_i \le -1$, and hence

$$(|R|+1)\,\mathbf{x}' \cdot \mathbf{t}_i + |R| \le (-|R| - 1) + |R| < 0 \ .$$

□

Lemma 42. *Let N be a net with a positive place invariant. Let $\{t_1, \ldots, t_k\}$ be a set of transitions of N, containing at most one transition of each conflict area. Then either the set $\{\mathbf{t}_1, \ldots, \mathbf{t}_k\}$ of associated columns of the incidence matrix is linearly independent or there are nonnegative coefficients $\lambda_1, \ldots, \lambda_k \ge 0$ such that $(\lambda_1, \ldots, \lambda_k) \ne \mathbf{0}$ and $\lambda_1 \mathbf{t}_1 + \cdots + \lambda_k \mathbf{t}_k = \mathbf{0}$.*

Proof. Assume the set $\{\mathbf{t}_1, \ldots, \mathbf{t}_k\}$ is not linearly independent. Then there exist coefficients $\mu_1, \ldots, \mu_k$ such that at least one of the μ_i is positive and

$$\mu_1 \mathbf{t}_1 + \cdots + \mu_k \mathbf{t}_k = \mathbf{0} \ .$$

For each i, $1 \le i \le k$, define the coefficient λ_i by

$$\lambda_i := \begin{cases} \mu_i & \text{if } \mu_i \ge 0, \\ 0 & \text{otherwise.} \end{cases}$$

No λ_i is negative. Since at least one of the μ_i is positive this likewise holds for some λ_i. Now define the transition vectors $\mathbf{j}$ and $\mathbf{j}'$ as follows:

$$\mathbf{j} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ and } \mathbf{j}' = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

The vector $\mathbf{j}$ is a transition invariant vector, i.e., $\mathbf{N}\cdot\mathbf{j} = \mathbf{0}$. We will prove $\mathbf{N}\cdot\mathbf{j}' \geq \mathbf{0}$.

Let s_i be any place and $\mathbf{s}_i$ its associated row vector of the incidence matrix. The negative components in $\mathbf{s}_i$ belong to the transitions in $s_i^\bullet$. In particular, they all belong to a single conflict area. Hence, by assumption, $s_i^\bullet$ contains at most one transition t in $\{t_1, \ldots, t_k\}$. For any other negative component of $\mathbf{s}_i$, the component in $\mathbf{j}$ and the component $\mathbf{j}'$ equals 0. We show $\mathbf{s}_i \cdot \mathbf{j}' \geq 0$.

If $s_i^\bullet \cap \{t_1, \ldots, t_k\} = \emptyset$ then nothing has to be shown. Now assume that $s_i^\bullet \cap \{t_1, \ldots, t_k\} = \{t\}$.

Case 1. If $\mathbf{j}(t) \leq 0$ then $\mathbf{j}'(t) = 0$ and we are finished.

Case 2. If $\mathbf{j}(t) > 0$ then $\mathbf{j}'(t) = \mathbf{j}(t)$. The vector $(\mathbf{j}' - \mathbf{j})$ has no negative components. Furthermore, $(\mathbf{j}' - \mathbf{j})(t) = 0$. So, for each negative component of $\mathbf{s}_i$, the component in the vector $(\mathbf{j}' - \mathbf{j})$ has the value 0. Therefore, $\mathbf{s}_i \cdot (\mathbf{j}' - \mathbf{j}) \geq 0$. Finally, $\mathbf{s}_i \cdot \mathbf{j} = 0$ implies $\mathbf{s}_i \cdot \mathbf{j}' \geq 0$.

So we obtain

$$\mathbf{N} \cdot \mathbf{j}' = \begin{pmatrix} \mathbf{s}_1 \\ \vdots \\ \mathbf{s}_n \end{pmatrix} \cdot \mathbf{j}' \geq \mathbf{0}.$$

The net N has a positive place invariant i, i.e. $\mathbf{i} > \mathbf{0}$ and $\mathbf{i} \cdot \mathbf{N} = \mathbf{0}$. Hence, $\mathbf{i} \cdot \mathbf{N} \cdot \mathbf{j}' = 0$. Now $\mathbf{i} > \mathbf{0}$ and $\mathbf{N} \cdot \mathbf{j}' \geq \mathbf{0}$ implies $\mathbf{N} \cdot \mathbf{j}' = \mathbf{0}$.

By definition of $\mathbf{j}'$, we finally obtain

$$\lambda_1 \mathbf{t}_1 + \cdots + \lambda_k \mathbf{t}_k = \mathbf{0}$$

which completes the proof. □

Theorem 43. *Let N be a net with positive place and transition invariants. If the rank of* $\mathbf{N}$ *is smaller than the number of conflict areas of N then there exists a live initial marking of N.*

Proof. The contraposition of this theorem is given by the implication (1) ⇒ (7) of the previously defined propositions (1) and (7). With the above lemmas, the implications (1) ⇒ (2) ⇒ $\cdots$ ⇒ (7) are shown above. □

This sufficient condition for potential liveness can be checked efficiently, i.e., in polynomial time: the existence of positive invariants can be decided by solving the respective inequality systems and the rank inequality is checked by calculating the rank and counting the conflict areas.

6.3 Characterizing Strongly Live Markings

In this subsection we give a sufficient condition for liveness of a marking. This condition is also necessary for strong liveness.

Lemma 44. *Let N be a net with positive place and transition invariants. Let $\{s_1, \ldots, s_k\}$ be a set of places of N containing at most one place of each conflict area. Then either the set of row vectors $\{\mathbf{s}_1, \ldots, \mathbf{s}_k\}$ of the incidence matrix is linearly independent or there exist nonnegative coefficients $\lambda_1, \ldots, \lambda_k \geq 0$ such that $(\lambda_1, \ldots, \lambda_k) \neq \mathbf{0}$ and $\lambda_1 \mathbf{s}_1 + \cdots + \lambda_k \mathbf{s}_k = \mathbf{0}$.*

Proof. Assume $N = (S, T, F)$. Let $N^d = (T, S, F^{-1})$ be the *dual net* of N, where $F^{-1} = \{(y, x) | (x, y) \in F\}$. Then $\mathbf{N}^d$ is the transposed matrix of $\mathbf{N}$. Each positive transition invariant of N is a positive place invariant of N^d, and vice versa. The net N is strongly connected because it has positive place and transition invariants (Theorem 9). In particular, no conflict area consists of a single transition with an empty pre-set. By definition, every conflict area of N^d is the pre-set (in N) of a conflict area of N, and vice versa. Therefore, the set $\{s_1, \ldots, s_k\}$ is a set of transitions of N^d, containing at most one transition of each conflict area. Now Lemma 42, applied to N^d and the set $\{s_1, \ldots, s_k\}$, proves the result. □

Theorem 45. *Let N be a net with positive place and transition invariants. Assume that the rank of $\mathbf{N}$ is smaller than the number of conflict areas of N. Then a marking m_0 of N is strongly live if and only if every non-negative place invariant vector $\mathbf{i} \geq \mathbf{0}$ satisfies $\mathbf{i} \cdot \mathbf{m}_0 > 0$.*

Proof.
($\Rightarrow$) Each strongly live marking is live. As shown in [DeRe98], a live marking m_0 satisfies $\mathbf{i} \cdot \mathbf{m}_0 > 0$ for each place invariant $\mathbf{i} \geq \mathbf{0}, \mathbf{i} \neq \mathbf{0}$.
($\Leftarrow$) We prove the contraposition. Let m_0 be a marking of N which is not strongly live. By Lemma 40, some marking m is reachable from m_0 such that no transition is strongly enabled at m. Hence, for each conflict area K, there exists at least one place in ${}^\bullet K$ that is not marked at m.

W.l.o.g. assume that $\{s_1, \ldots, s_k\}$ is a set of places, all unmarked at m, containing for each conflict area K exactly one place of ${}^\bullet K$. The number of places in this set is equal to the number of conflict areas. By assumption, the rank of $\mathbf{N}$ is smaller than the number of conflict areas. Hence the set of vectors $\{\mathbf{s}_1, \ldots, \mathbf{s}_k\}$ is not linearly independent. By Lemma 44, there are coefficients $\lambda_1, \ldots, \lambda_k \geq 0$ such that $(\lambda_1, \ldots, \lambda_k) \neq \mathbf{0}$ and $\lambda_1 \mathbf{s}_1 + \cdots + \lambda_k \mathbf{s}_k = \mathbf{0}$.

Let $\mathbf{i}$ be the place vector of N, defined by

$$\mathbf{i} = (\lambda_1, \ldots, \lambda_k, 0 \ldots, 0) .$$

Then $\mathbf{i} \cdot \mathbf{N} = \mathbf{0}$, i.e., $\mathbf{i}$ is a place invariant vector satisfying $\mathbf{i} \geq \mathbf{0}, \mathbf{i} \neq \mathbf{0}$. Since every place in $\{s_1, \ldots, s_k\}$ is unmarked at m, we have $\mathbf{i} \cdot \mathbf{m} = 0$. Since m is reachable from m_0 and since $\mathbf{i}$ is a place invariant, finally follows $\mathbf{i} \cdot \mathbf{m}_0 = 0$. □

The following corollary formulates the sufficient condition for liveness of a marked net given by Theorems 43 and 45.

Corollary 46. *Let N be a net with initial marking m_0. Assume that N has positive place and transition invariants and that the rank of* $\mathbf{N}$ *is smaller than the number of conflict areas of N. If every place invariant* $\mathbf{i}$ *with* $\mathbf{i} \geq 0, \mathbf{i} \neq 0$ *satisfies* $\mathbf{i} \cdot \mathbf{m}_0 > 0$ *then m_0 is a live marking.*

The condition of the previous theorem and corollary can be checked efficiently by help of the following system of inequalities. The condition is satisfied if and only if there exists no solution

$$\begin{array}{rl} \mathbf{N}^{\mathsf{T}} \cdot \mathbf{x} = 0 & (\mathbf{x}^{\mathsf{T}} \text{ is a place invariant vector}) \\ (1,0,\ldots,0) \cdot \mathbf{x} \geq 0 & (\mathbf{x} \geq 0) \\ \vdots & \\ (0,\ldots,0,1) \cdot \mathbf{x} \geq 0 & \\ (1\ldots1) \cdot \mathbf{x} \geq 1 & (\mathbf{x} \neq 0) \\ \mathbf{m}^{\mathsf{T}} \cdot x \leq 0\,. & \end{array}$$

6.4 A Necessary Condition for the Liveness of a Marking

Now we show the counterpart to the previous results, i.e. we provide a necessary condition for liveness which is based on the rank of the incidence matrix, too.

As motivated at the beginning of this section, two transitions which are always enabled in conflict lead to different transition invariants. This observation holds in particular for transitions with identical pre-sets. In the sequel, we add *regulation circles* to nets, that ensure that sets of transitions with identical pre-sets occur cyclically in a fixed order.

Let $U = \{t_1, \ldots, t_k\}$ be a nonempty set of transitions of a net. For each transition $t_i \in U$, let s'_i be a new place. (s'_i is not an element of N). Then the net

$$N^U = (\{s'_1, \ldots, s'_k\}, U, \{(t_1, s'_1), (s'_1, t_2), \ldots, (t_k, s'_k), (s'_k, t_1)\})$$

is called a *regulation circle* of U. The net obtained by the componentwise union of the nets N and N^U is called the *composition* of N and N^U.

Lemma 47. *Let N be a net, let N^U be the regulation circle of some set U of transitions of N, and let N' be the composition of N and N^U. If N has a positive place invariant then N' also has a positive place invariant.*

Proof. Let $\mathbf{i}$ be a positive place invariant vector of N. Let $\mathbf{i}'$ be a place vector of N' that coincides with $\mathbf{i}$ for all places of N and that has the value 1 for all places of the regulation circle. By construction, this vector is positive.

Let t be an arbitrary transition of N. If $t \notin U$ then ${}^\bullet t$ in N and ${}^\bullet t$ in N' are identical, and the same holds for $t^\bullet$. If $t \in U$, both ${}^\bullet t$ and $t^\bullet$ contain in N' exactly one additional place of the regulation circle. Both corresponding components of $\mathbf{i}'$ are 1. So, in both cases, $\mathbf{i}' \cdot \mathbf{t} = \mathbf{i} \cdot \mathbf{t} = 0$. □

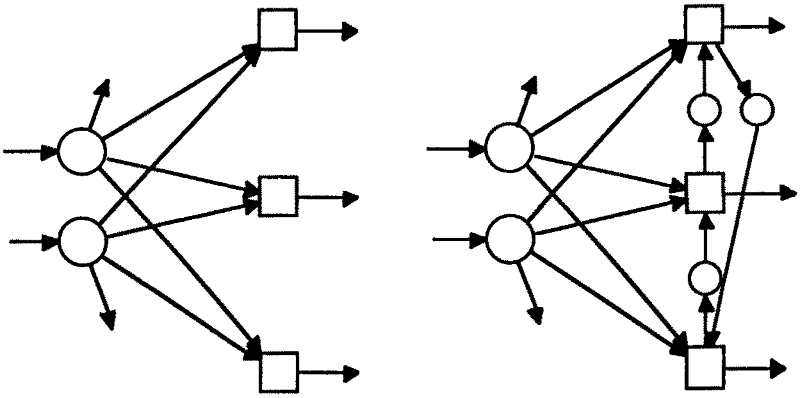


Fig. 8. A piece of a net with a regulation circle

Lemma 48. *Let N be a net with a positive place invariant. Let U be a nonempty set of transitions of N such that all transitions in U have identical pre-sets. Let N^U be a regulation circle of U, and let N' be the composition of N and N^U. If there exists a live marking of N then there also exists a live marking of N'.*

Proof. Let m_0 be a live marking of N. Since N has a positive place invariant , m_0 is bounded. Therefore, the set of markings reachable from m_0 is finite; let k be its cardinality. Define a marking m_0' of the net N' that coincides with m_0 for all places of N and associates k tokens to every place of the regulation circle. We will show that m_0' is a live marking of N'.

Let $m_0' \xrightarrow{\sigma} m_1'$ be any occurrence sequence of N', and t be any transition. We will construct an occurrence sequence τ, enabled at m_1', that includes an occurrence of t. The sequence τ is the concatenation of two sequences τ_1 and τ_2.

Since N' is constructed from N and N^U by identification of common transitions only, every occurrence sequence σ of N' induces occurrence sequences of N and N^U. More formally, for each marking m' of N' we denote its restriction to the places of N by m and its restriction to the places of N^U by m^U. The projection σ^U of σ to the transitions of N^U is obtained from σ by cancellation of all transitions that are not in U. Now it is easy to verify that there is an occurrence sequence

$$m_0' \xrightarrow{\sigma} m_1' \text{ of } N'$$

if and only if there are occurrence sequences

$$m_0 \xrightarrow{\sigma} m_1 \text{ of } N \text{ and}$$

$$m_0^U \xrightarrow{\sigma^U} m_1^U \text{ of } N^U \quad .$$

(Notice that N and N' have identical sets of transitions.)

The net N^U is a simple circle. Therefore, its initial marking m_0^U (associating k tokens to every place) can be reached again from m_1^U. Let

$$m_1^U \xrightarrow{u_1 \dots u_n} m_0^U$$

be an according occurrence sequence of the net N^U.

We show that the marking m_1 of the net N enables an occurrence sequence τ_1 such that its projection ${\tau_1}^U$ is $u_1 \dots u_n$. To this end, we choose a minimal occurrence sequence enabled at m_1 such that the reached marking enables a transition in U (this might be the empty sequence). Since m_0 is live, m_1 is live, too. So, such an occurrence sequence exists. Since all transitions in U have identical pre-sets in N, they all are enabled at m_1. Now extend the occurrence sequence by the transition u_1. Starting with the marking reached after u_1, choose again a minimal occurrence sequence enabling the transitions of U, and then continue with u_2. By the minimality of the occurrence sequences, these sequences do not contain any transition of U. After n repetitions of this procedure, an occurrence sequence $m_1 \xrightarrow{\tau_1} m_2$ is gained such that the projection of τ_1 to the set U is the sequence $u_1 \dots u_n$.

The sequence τ_1 is enabled at m_1 in N and leads to m_2. Its restriction $\tau_1^U = u_1 \dots u_n$ is an occurrence sequence of N_U leading from m_1^U to m_0^U. Hence the occurrence sequence τ_1 is also enabled at m'_1 in the net N'. It leads to the marking m'_2 which coincides with m_2 on the places of N and with m'_0 on the places of N^U.

The marking m_2 is again a live marking of N. So, m_2 enables a minimal occurrence sequence τ_2 that contains the transition t. By minimality, no marking is reached more than once during this occurrence sequence. So the length of τ_2 is smaller than k, the number of all reachable markings. Since, in N^U, the marking m_0^U associates k tokens to each place, it enables all sequences with length up to k and in particular the sequence τ_2^U. Therefore, the sequence τ_2 is enabled at m'_2 in N'. Finally, the composition $\tau_1\, \tau_2$ is enabled at m'_0 and contains t. □

For a net N, the set $P_N = \{{}^\bullet t \mid t \in T_N\}$ denotes all pre-sets of transitions of N (T_N is the set of transitions of N). If two transitions have identical pre-sets, they only contribute to the set P_N once.

Theorem 49. *Let N be a net with positive place and transition invariants. If N has a live marking then the rank of* **N** *is smaller than the value $|P_N|$.*

Proof. Assume that N has a live marking m_0. We proceed by induction on $k = |T_N| - |P_N|$.

Base. Assume $k = 0$. The matrix **N** has $|T_N|$ columns. The rank of **N** is at most $|T_N| - 1$ because N has a positive transition invariant. Since $k = 0$ we have $|T_N| = |P_N|$. So the rank of **N** is smaller than $|P_N|$

Step. Assume $k > 0$. Let U be a maximal set of transitions with identical pre-sets. We have $|U| \geq 2$ because $k = |T_N| - |P_N| > 0$.

Let N^U be the regulation circle of U and let N' be the composition of N and N^U. By Lemmas 47 and 48, N' has a positive place invariant and can be marked lively. So it has a live and bounded marking. As shown in [DeRe98], it has a positive transition invariant.

The set of transitions of N' is also T_N. Each transition in $T_N \setminus U$ has the same pre-set in N and in N'. In N', two distinct transitions of U have different pre-sets, since this is the case in the regulation circle. So we get

$$|P_{N'}| = |P_N| + |U| - 1 \ .$$

This implies that the number $|T_N| - |P_{N'}|$ is smaller than $|T_N| - |P_N|$. Hence, with the induction hypotheses applied to the net N', the rank of $\mathbf{N'}$ does not exceed $|P_{N'}| - 1$. So we have

$$\text{rank}(\mathbf{N}) + |U| - 1 \leq \text{rank}(\mathbf{N'}) \leq |P_{N'}| - 1 = |P_N| + |U| - 2 \ .$$

It remains to show

$$\text{rank}(\mathbf{N}) + |U| - 1 \leq \text{rank}(\mathbf{N'}),$$

i.e. the rank of $\mathbf{N'}$ exceeds the rank of $\mathbf{N}$ by at least $|U| - 1$. W.l.o.g. assume $U = \{t_1, \ldots, t_n\}$. The n rows in $\mathbf{N'}$ of the places of the regulation circle are shaped:

$$\begin{matrix} (\, -1 & 0 \cdots & 0 & 1\, 0 \cdots 0\,) \\ (\;\; 1 & -1 & 0 \cdots & 0\, 0 \cdots 0\,) \\ \vdots & \ddots & & \vdots\;\vdots \quad \vdots\;\vdots \\ (\;\; 0 \cdots & & 1 \; -1 & 0\, 0 \cdots 0\,) \\ (\;\; 0 \cdots & & 0 & 1 \, -1\, 0 \cdots 0\,) \end{matrix}$$

Clearly, every subset of this set containing $n-1$ vectors is linearly independent. We show that none of this rows is a linear combination of other rows of $\mathbf{N'}$, i.e. of rows of $\mathbf{N}$. We proceed indirectly and assume that the row $\mathbf{s}'_i$ is a linear combination of the rows of $\mathbf{N}$. Let $\mathbf{s}'_i = \mathbf{k} \cdot \mathbf{N}$, let t_i be the unique transition in the pre-set of s'_i, and let t_j be the unique transition in the post-set of s'_i. By assumption, there exists a live marking m_0 of N. Since N has a positive place invariant, m_0 is a bounded marking and only finitely many markings can be reached. Let m be reachable from m_0 such that $\mathbf{k} \cdot \mathbf{m}$ has a maximal value. Define

$$\lambda := \mathbf{k} \cdot \mathbf{m} - \mathbf{k} \cdot \mathbf{m}_0 \ .$$

Since, in N, the transitions t_i and t_j have identical pre-sets, there exists an occurrence sequence $m_0 \xrightarrow{\sigma} m_1$ such that the transition t_i occurs $\lambda + 1$ times in σ and t_j does not occur in σ. The corresponding Parikh vector $\mathbf{p}_\sigma$ satisfies $\mathbf{s}'_i \cdot \mathbf{p}_\sigma > \lambda$. By $\mathbf{m}_0 + \mathbf{N} \cdot \mathbf{p}_\sigma = \mathbf{m}_1$ and by the choice of m follows

$$\mathbf{k} \cdot \mathbf{N} \cdot \mathbf{p}_\sigma = \mathbf{k} \cdot \mathbf{m}_1 - \mathbf{k} \cdot \mathbf{m}_0 \leq \lambda,$$

in contradiction to

$$\mathbf{k} \cdot \mathbf{N} \cdot \mathbf{p}_\sigma = \mathbf{s}'_i \cdot \mathbf{p}_\sigma > \lambda.$$

□

Bibliographic Remarks

The rank conditions were first introduced for *free-choice nets*. There the sufficient and the necessary condition coincide.

The rank conditions for free-choice nets were developed in [Espa90], [CaCS91] and [Dese92] (see also [DeEs95]). They were generalized to free-choice nets with arc weights (*equal-conflict nets*) in [ReTS95, TeSi96].

The necessary rank condition for general nets is from [CoCS90]; see also [DeEs95] and [TeSi96]. The sufficient condition is from [Dese94].

In [Reis79], conflict areas were introduced.

The paper [Mura77] provides further bounds for the rank of the incidence matrix: If each marking of a net is reachable from some initial marking then the row rank is maximal and equals the number of places. A net with this property has no place invariant except the vector **0**.

References

[CaCS91] G. Campos, G. Chiola and M. Silva. Properties and performance bounds for closed free choice synchronized monoclass queueing networks. IEEE Transactions on Automation and Control, 36(12):1368–1382, 1991.

[CoCS90] J. M. Colom, J. Campos and M. Silva. On Liveness Analysis Through Linear Algebraic Techniques. Dpto. Ing. Electrica e Informatica, Universidad de Zaragoza RR 90-11, 1990.

[CoSi91] J. M. Colom and M. Silva. Improving the linearly based characterization of P/T nets. In G. Rozenberg (ed.), Advances in Petri nets 1990, Vol. 483 of Lecture Notes in Computer Science, pp. 113–145, Springer-Verlag, 1991.

[DeEs93] J. Desel and J. Esparza. Reachability in cyclic extended free choice systems. Theoretical Computer Science, 114:93–118, 1993.

[DeEs95] J. Desel and J. Esparza. Free Choice Petri Nets, Vol. 40 of Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1995.

[DeNR96] J. Desel, K.-P. Neuendorf and M.-D. Radola. Proving non-reachability by modulo-invariants. Theoretical Computer Science, 153:49–64, 1996.

[Dese85] J. Desel. Another boatman story. GI Petri Net Newsletter, 22:3–6, 1985.

[Dese88] J. Desel. Wanted: Dead or alive? GI Petri Net Newsletter, 29:0–2, 1988.

[Dese92] J. Desel. A proof of the rank theorem for extended free choice nets. In K. Jensen (ed.), Application and Theory of Petri nets, Vol. 616 of Lecture Notes in Computer Science, pp. 134–153, Springer-Verlag, 1992.

[Dese94] J. Desel. Regular marked Petri nets. In J. van Leeuwen (ed.), Graph-Theoretic Concepts in Computer Science, Vol. 790 of Lecture Notes in Computer Science, pp. 264–275, Springer-Verlag, 1994.

[DeRe98] J. Desel and Wolfgang Reisig. Place/transition Petri nets. Advanved Course on Petri Nets, 1996, in this volume.

[Dese98] J. Desel. Petrinetze, Lineare Algebra und Lineare Programmierung. Teubner-Texte zur Informatik Vol. 26, Teubner Verlag, 1998.

[EsMe96] J. Esparza and S. Melzer. Checking system properties via Integer Programming. In H.R. Nielson (ed.), ESOP'96, Vol. 1058 of Lecture Notes in Computer Science, pp. 250-264, Springer-Verlag, 1996.

[EsNi94] J. Esparza and M. Nielsen. Decidability issues for Petri nets. GI Petri Net Newsletter, 47:5–23, 1994.

[Espa90] J. Esparza. Synthesis rules for Petri nets, and how they lead to new results. In J.C.M. Baeten and J.W. Klop (eds.), CONCUR 90, Vol. 458 of Lecture Notes in Computer Science, pp. 182–198. Springer-Verlag, 1990.

[EzCS93] J. Ezpeleta, J. M. Couvreur and M. Silva. A new technique for finding a generating family of siphons, traps and ST-components. Application to colored Petri nets. In G. Rozenberg (ed.), Advances in Petri nets 1993, Vol. 674 of Lecture Notes in Computer Science, pp. 126–147, Springer-Verlag, 1993.

[Jant87] M. Jantzen. Complexity of place/transition nets. In W. Brauer, W. Reisig and G. Rozenberg (eds.), Petri nets: Central Models and Their Properties, Advances in Petri Nets 1986, part I, Vol. 254 of Lecture Notes in Computer Science, pp. 413–435, Springer-Verlag, 1987.

[JaVa80] M. Jantzen and R. Valk. Formal properties of place/transition nets. In W. Brauer (ed.), Net Theory and Applications, Vol. 84 of Lecture Notes in Computer Science, pp. 165–212, Springer-Verlag, 1980.

[KaBa79] R. Kannan and A. Bachem. Polynomial algorithms for computing the Smith and Hermite normal forms of an integer matrix. SIAM Journal on Computing, 8:499–507, 1979.

[Laut75] K. Lautenbach. Liveness in Petri nets. GMD-ISF 75-02.1, Gesellschaft für Mathematik und Datenverarbeitung, Bonn, 1975.

[Laut87a] K. Lautenbach. Linear algebraic techniques for place/transition nets. In W. Brauer, W. Reisig and G. Rozenberg (eds.), Petri nets: Central Models and their Properties, Advances in Petri nets 1986, Part I, Vol. 254 of Lecture Notes in Computer Science, pp. 142–167. Springer-Verlag, 1987.

[Laut87b] K. Lautenbach. Linear algebraic calculation of deadlocks and traps. In H.J. Genrich, K. Voss and G. Rozenberg (Ed.), Concurrency and Nets, pp. 315–336, Springer-Verlag, 1987.

[LaSc74] K. Lautenbach and H. A. Schmidt. Use of Petri nets for proving correctness of concurrent process systems. In IFIP Congress 1974, pp. 187–191. North-Holland, 1974.

[Lien76] Y. E. Lien. Termination properties of generalized Petri nets. SIAM Journal on Computing, 5(2):251–265, 1976.

[MeRo80] G. Memmi and G. Roucairol. Linear algebra in net theory. In W. Brauer (ed.), Net Theory and Applications, Vol. 84 of Lecture Notes in Computer Science, pp. 213–223. Springer-Verlag, 1980.

[Mura77] T. Murata. State equation, controllability, and maximal matchings of Petri nets. IEEE Transactions on Automation and Control, 22(3):412–416, 1977.

[Mura89] T. Murata. Petri nets: Properties, analysis and applications. Proc. of the IEEE, 77(4):541–580, 1989.

[Pasc86] K.-H. Pascoletti. Diophantische Systeme und Lösungsmethoden zur Bestimmung aller Invarianten in Petri-Netzen. GMD-Berichte Nr. 160, Oldenbourg, 1986.

[Pete77] J. L. Peterson. Petri nets. ACM Computing Surveys, 9(3):223–252, 1977.

[Pete81] J. L. Peterson. Petri Net Theory and the Modelling of Systems. Prentice-Hall, Englewood Cliffs, 1981.

[Petr73] C. A. Petri. Concepts of net theory. In: Mathematical Foundations of Computer Science, Proceedings of a symposium and summer school, High

Tatras, September 3-8, 1973. Math. Inst. of the Slovak Acad. of Science, pp. 137–146, 1973.

[Petr82] C. A. Petri. State-transition structure in physics and computation. International Journal of Theoretical Physics, 21(10/11):979-992, 1982.

[Ramc74] C. Ramchandani. Analysis of asynchronous concurrent systems by Petri nets. PhD thesis, MIT, Dept. Electrical Engineering, Cambridge, Mass., 1974.

[Reis79] W. Reisig. On solving conflicts in Petri nets. In: U. Pape (ed.),Discrete Structures and Algorithms, Hanser-Verlag, pp. 241–254, 1979.

[Reis85] W. Reisig. Petri Nets. EATCS Monographs on Theoretical Computer Science Vol. 4, Springer-Verlag, 1985.

[ReTS95] L. Recalde, E. Teruel and M. Silva. On well-formed analysis: The case of deterministic systems of sequential processes. In J. Desel (ed.), Structures in Concurrency Theory, Workshops in Computing, pp. 279–293, Springer-Verlag, 1995.

[Schr86] A. Schrijver. Theory of Linear and Integer Programming. Wiley, 1986.

[Sifa79] J. Sifakis. Structural properties of Petri nets. In J. Winkowski (ed.), Mathematical Foundations of Computer Science, Vol. 64 of Lecture Notes in Computer Science, pp. 474–483, Springer-Verlag, 1979.

[TeCS93] E. Teruel, M. Colom and M. Silva. Linear analysis of deadlock-freeness of Petri net models. 2nd European Control Conference, Vol. 2, pp. 513–518. North–Holland, 1993.

[TeSi96] E. Teruel and M. Silva. Structure theory of equal conflict systems. Theoretical Computer Science, 153:271–300, 1996.