Place/Transition Petri Nets

Jörg Desel[1] and Wolfgang Reisig[2]

[1] Institut für Angewandte Informatik und Formale Beschreibungsverfahren
Universität Karlsruhe, Germany
desel@aifb.uni-karlsruhe.de
[2] Institut für Informatik
Humboldt-Universität zu Berlin, Germany
reisig@informatik.hu-berlin.de

Abstract. This contributions provides an introduction to the theory of place/transition Petri nets. Topics include the sequential and the concurrent behavior of place/transition Petri nets, marking graphs and coverability trees, and some analysis techniques that are based on the structure of place/transition Petri nets.

1 Introduction

Place/transition Petri nets (p/t-nets, for short) are the most prominent and best studied class of Petri nets. This class is sometimes just called *Petri nets*. In contrast to *elementary (Petri) net systems* [RoEn98], in p/t-nets a place can hold any number of tokens. In contrast to *high-level Petri nets* [Jens98], these tokens are indistinguishable. This contribution surveys some of the the most important concepts and analysis techniques for p/t-nets. The reader is assumed to have a rough idea about Petri nets, and in particular about elementary net systems. However, this knowledge is not a prerequisite for this paper.

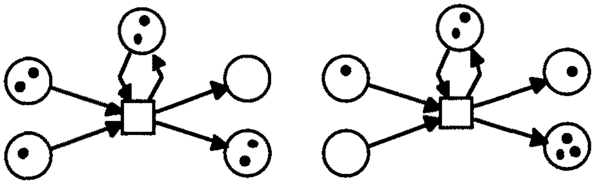
A state of a p/t-net is determined by a distribution of tokens on its places. It is formally defined as a mapping that assigns each place a nonnegative integer. This mapping is usually called a *marking*. A p/t-net equipped with an initial marking constitutes a Petri net system[3], i.e. a system model capturing both static structural and dynamic behavioral aspects.

This paper mainly concentrates on a basic class of marked p/t-nets, which only consist of a net and an initial marking. These Petri net systems are called *marked nets*. We only roughly survey some additional features of p/t-nets. The basic concepts of arbitrary p/t-nets can be presented more lucidly for marked nets. Most of them can easily be generalized to more general classes of p/t-nets.

The elementary behavior of a marked net is defined by a local occurrence rule that determines the possible elementary marking changes caused by transition occurrences. A transition of a marked net is enabled at a marking if each arc to the transition originates at a place which carries at least one token. The

[3] sometimes called place/transition-system

occurrence of an enabled transition is described by means of the following picture, representing a transition and its vicinity before and after its occurrence:



Starting with the initial marking, transition occurrences yield new markings, that in turn give rise to further transition occurrences. A run of a marked net consists of a set of transition occurrences that altogether might happen in some order and hence do not exclude each other. The global behavior of a marked net is determined by its runs.

There are various ways to formalize runs. We will emphasize a sequential view, given by sequences of transition occurrences (*occurrence sequences*), and a concurrent view, given by partially ordered *process nets*. Sequential runs are technically simple. However, concurrent occurrences of transitions have to be ordered in sequential runs whereas concurrency and dependency is explicitly represented in concurrent runs.

Occurrence sequences will be used to define behavioral properties of marked nets, such as *deadlock-freedom* (no total deadlock is reachable), *liveness* (no partial deadlock is reachable), *boundedness* (on each place, the token load does not grow unlimited), and *reversibility* (the initial marking can always be reached again). Analysis of a marked net w.r.t. these properties will be based on finite graphs that comprise the sequential behavior of a marked net, namely *marking graphs* and *coverability trees*. It will also be demonstrated that some of the properties can be characterized in terms of process nets, and that this view allows more insight to behavior and to more elegant proofs.

This contribution compiles fundamentals of theoretical results about p/t-nets. Some pointers to literature should provide a starting point for further work. For textbooks on Petri nets in general see [Pete81, Reis85]. Another good introductory text is [Mura89]. A collection of surveys is given in [BRR87]. There, the papers [Reis87, Laut87] are particularly related to p/t-nets.

Decidability and complexity results on p/t-nets are important topics, though missing in this contribution. The most important result in this field is the decidability of the problem whether a given marking is reachable in a marked p/t-net. A solution to this *reachability problem* was given in [Mayr84, Kosa82], see also [Reut88]. The paper [Jant87], also contained in [BRR87], gives in overview on some more results in this area. More recent papers surveying in particular results on *model checking* and p/t-nets are [EsNi94], [Espa98] and [Valm98], the latter two in this volume.

A marking of a p/t-net can be conceived as a tuple of numbers, each of which describing the number of tokens residing at some place. This view suggests a marking to be conceived as a vector of nonnegative integers. The change of a marking, caused by the occurrence of a transition corresponds to the addition of an integer-valued vector. Therefore, p/t-nets are isomorphic to a class of vector addition systems. In fact, a lot of Petri net research has originally focussed on vector addition systems. The survey paper [Dese98] in this volume presents application of linear algebraic techniques to the analysis and verification of marked nets.

A further research area focusses on structurally defined subclasses of Petri nets. The best known examples of such classes are *marked graphs*, *state machines* and *free choice nets*. In a marked graph, no place is branching. Therefore, marked graphs behave deterministically and do not allow to model *choice* between actions. A state machine has no branching transitions and hence does not allow for *synchronization* of actions. In free-choice nets, the combination of forward branched places and backward branched transitions is restricted, thus ruling out a too close interplay between choice and synchronization. Free-choice nets generalize the classes mentioned above. As pointed out in the present paper and in [Dese98], efficient analysis and verification techniques for arbitrary nets stick to either sufficient or necessary conditions for behavioral properties, whereas decision algorithms for behavioral properties suffer from the huge inherent complexity of most problems. For free-choice nets, the picture looks significantly better because many interesting behavioral properties of marked free-choice nets can be characterized in terms of the net structure. For a detailed treatment of the theory of free-choice Petri nets, the authors warmly recommend the textbook [DeEs95]. This book also contains many references to the original sources.

The following section provides an introductory example. In the third section, marked nets will be defined formally. We present the *Strongly-Connectedness Theorem* as an example for relations between structure and behavior of a marked net. The fourth section is devoted to *marking graphs* and *coverability trees* and to the representation of behavioral properties of marked nets by means of structural properties of these graphs. In section five, more efficient analysis techniques based on the structure of nets are presented, namely *place invariants*, *transition invariants*, *siphons*, and *traps*. Advantages of process nets representing the concurrent runs of marked nets as well as their relation to occurrence sequences are presented in section six. Finally, the seventh section provides a quick survey on additional features of p/t-nets, including *arc weights*, *capacity restrictions* and *inhibitor arcs*.

2 An Introductory Example

Marked p/t-nets generalize *contact-free elementary net systems* [RoEn98]. In other words: every contact-free elementary net system is at the same time a marked p/t-net[4]. We start with an example of an elementary net system and subsequently add features of p/t-nets which are not available at the elementary level.

[4] Elementary net systems can nevertheless be considered *elementary* since all basic concepts of p/t-nets can be derived from respective concepts of elementary net systems, see [Golt90, DeMe90].

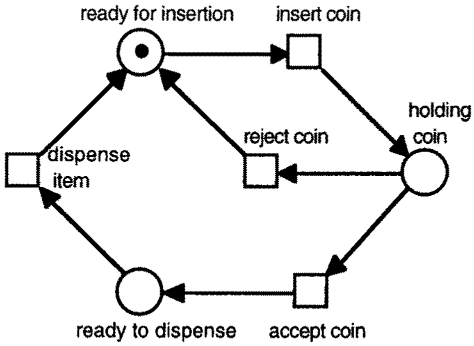


Fig. 1. A marked net representing the control structure of a vending machine

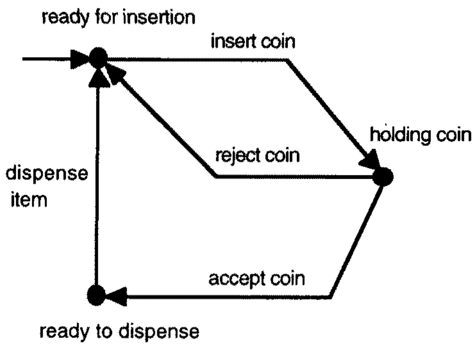


Fig. 2. The marking graph of the above marked net

Figure 1 shows an elementary net system modeling the control part of a vending machine. At the initial state shown in the figure, the machine is waiting for a coin to be inserted. This is modeled by the token on the place `ready for insertion`. An inserted coin is either rejected or accepted, depending on the result of a validity check which is not part of the system model. If the coin was rejected, the system returns to its initial state. Otherwise the system first dispenses an item and then returns to its initial state.

This system model is also a marked p/t-net. Terminology for p/t-nets usually differs from terminology for elementary net systems. Nonetheless we shall introduce and use the place/transition based terminology already here:

- every condition of an elementary net system is a *place* of the p/t-net,
- every event of an elementary net system is a *transition* of the p/t-net,
- every state or case of an elementary net system is a *marking* of the p/t-net; formally, a marking is represented as a mapping assigning each place, the number of tokens it carries – here 0 or 1.

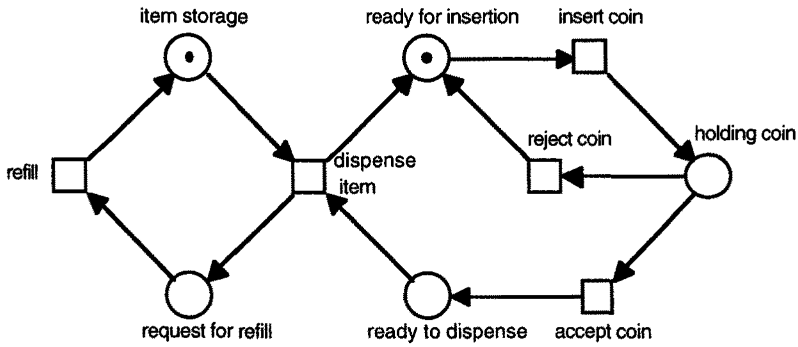


Fig. 3. The vending machine with capacity 1

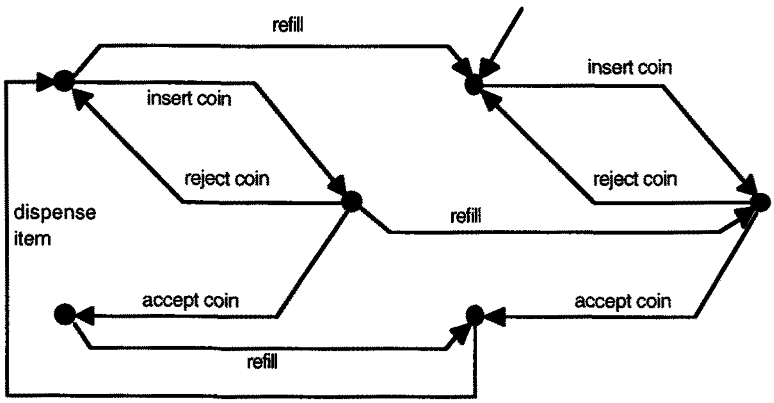


Fig. 4. The marking graph of the above marked net

The sequential behavior of a p/t-net can be represented by its *marking graph*. In case there is always only one token in the net, as in the above example, the marking graph is quite small, and it resembles the net itself (see Figure 2). In general, the marking graph is a compact representation of all possible sequences of subsequent transition occurrences. Starting with the distinguished *initial vertex* (`ready for insertion` in the example), each directed path of the marking graph represents a possible sequence of transition occurrences, and vice versa.

Now we add the system component delivering items to the vending machine. Assume the capacity of the machine is 1, i.e., after delivering one item, the machine has to be refilled (by another component not considered here), as shown in Figure 3. This system is still an elementary net system. Its marking graph is shown in Figure 4. For sake of readability, markings annotating vertices are dropped. The shape of the graph should help to compare it with the graph of Figure 2.

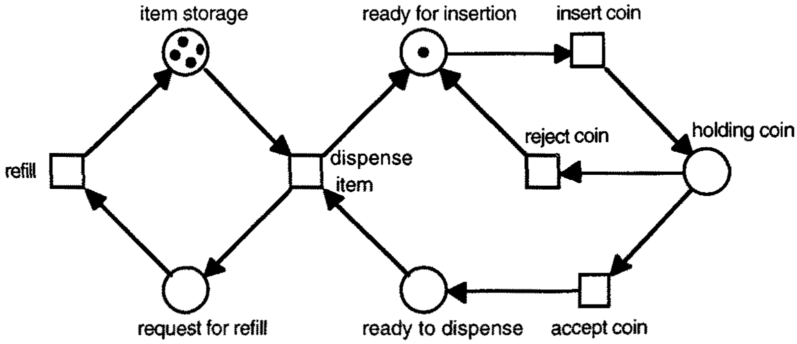


Fig. 5. The vending machine with capacity 4

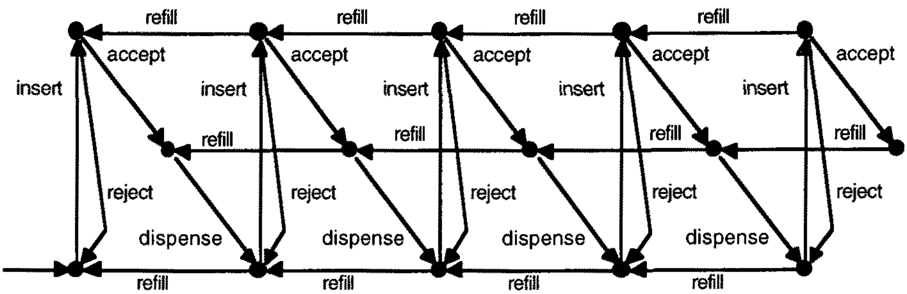


Fig. 6. The marking graph of the above marked net

The marking graph suggests that the system exhibits no concurrency. After the occurrences of transitions **`insert coin`**, **`accept coin`** and **`dispense item`**, both transitions **`refill`** and **`accept coin`** are enabled. The corresponding marking corresponds to the upper left hand vertex in the graph. Although the transitions **`refill`** and **`accept coin`** can occur concurrently, the sequential view enforces an order between these two transition occurrences. We will come back to this example and discuss its concurrent semantics in section six.

The storage capacity of real vending machines exceeds 1, of course. Figure 5 shows the same net as the previous example. However, the initial marking of the net is different: the place **`item storage`** carries four tokens initially. This systems is no elementary net system anymore: the place **`item storage`** cannot be interpreted as a condition with two states *true* and *false*. Instead, it has five possible states:

unmarked, marked with one token, ..., marked with four tokens.

The transition **`dispense item`** is only enabled if each of the two places **`ready to dispense`** and **`item storage`** contains at least one token. An occurrence of **`dispense item`** removes a token from **`ready to dispense`** and a token from

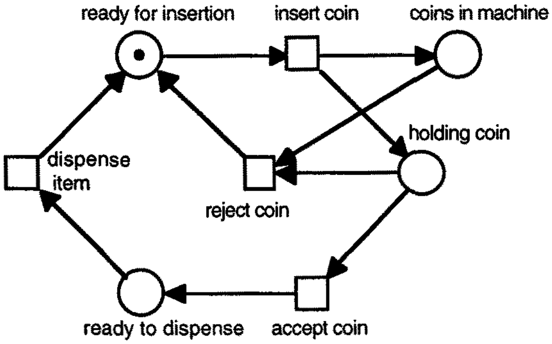


Fig. 7. The control structure of a vending machine with a counter

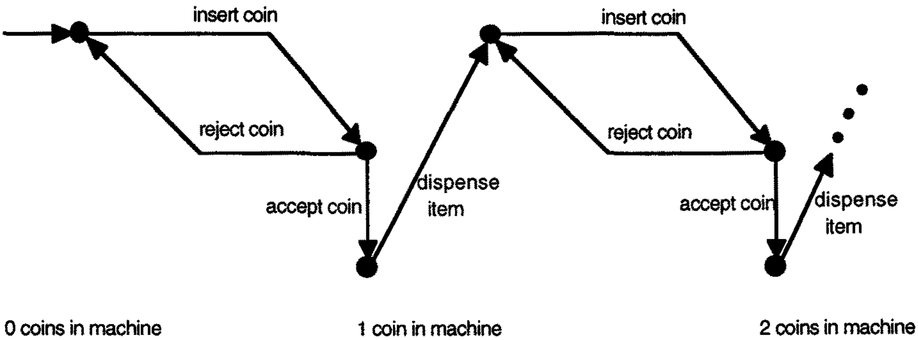


Fig. 8. The marking graph of the above marked net

`item storage`, and adds tokens to both places `request for refill` and `ready for insertion`.

All places of the marked net shown in Figure 5 are *bounded*: every place possesses an upper bound for the number of tokens it will carry at any reachable marking. This is not the case for the marked net shown in Figure 7. For simplicity, the left part of the previous net is omitted here, and an additional place represents the number of coins that are currently in the machine. Since this number can grow without limit, this place is *unbounded*. It can take an infinite number of different states, whence the marking graph of the marked net is infinite (see Figure 8). We will show in section 4 how finite *coverability trees* are used for analysis of unbounded marked nets.

As a further property of marked nets, the notion of *liveness* will intensely be studied in the sequel. A marked net is live if every transition can always occur again; more precisely, if from any reachable marking it is possible to reach some marking that enables the transition. All marked nets modeling the vending machine considered so far are live.

The property of *deadlock-freedom* is weaker than liveness: a marked net is deadlock-free if every reachable marking enables some transition.

A marked net is *reversible* if it can always reach its initial marking again. The marked nets of Figures 1, 3 and 5 are reversible, whereas the net of Figure 7 is not.

3 Basic Definitions and Elementary Results

This section formally defines marked nets and their behavior. As an example for the interplay between statical properties of nets and dynamic properties of marked nets, the Strongly-Connectedness-Theorem is proven.

3.1 Petri Nets

A *net* N is constituted by

- a set S of *places*,
- a set T of *transitions* such that $S \cap T = \emptyset$, and
- a set F of *directed arcs* (*flow relation*), $F \subseteq (S \cup T) \times (S \cup T)$, satisfying

$$F \cap (S \times S) = F \cap (T \times T) = \emptyset.$$

All places and transitions are said to be *elements* of N. A net is *finite* if its set of elements is finite.

In the graphical representation of a net, places are drawn as circles, transitions as squares and arcs as arrows. An arrow with two arrowheads between elements x and y indicates that (x, y) as well as (y, x) are directed arcs.

The sets of places, transitions and arcs of a net N are denoted by S_N, T_N and F_N, respectively. The set of all elements of the net N is sometimes also denoted by N, i.e., $N = S_N \cup T_N$.

For an element x of N, its *pre-set* ${}^\bullet x$ is defined by

$${}^\bullet x = \{y \in N \mid (y, x) \in F_N\}$$

and its *post-set* $x^\bullet$ is defined by

$$x^\bullet = \{y \in N \mid (x, y) \in F_N\}.$$

For $X \subseteq N$, define

$${}^\bullet X = \bigcup_{x \in X} {}^\bullet x \quad \text{and} \quad X^\bullet = \bigcup_{x \in X} x^\bullet.$$

An element x is *isolated* if ${}^\bullet x = x^\bullet = \emptyset$.

A *directed path* (*path* for short) of a net is a nonempty sequence $x_0 \ldots x_k$ of elements satisfying $x_i \in x_{i-1}^\bullet$ for each i $(1 \leq i \leq k)$. This path *leads from* x_0 *to* x_k. The net is *strongly connected* if for each two elements x and y there exists a directed path leading from x to y.

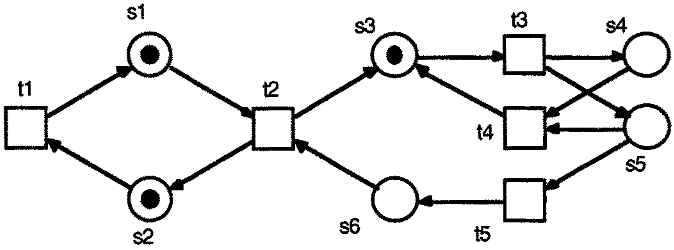


Fig. 9. A marked Petri net

An *undirected path* is a nonempty sequence $x_0 \dots x_k$ of elements satisfying $x_i \in {}^\bullet x_{i-1} \cup x_{i-1}^\bullet$ for each i $(1 \leq i \leq k)$. It *leads from* x_0 *to* x_k. The net is *weakly connected* if, for each two elements x and y, there exists an undirected path leading from x to y.

In Figure 9, a net is shown with places {s1, s2, s3, s4, s5, s6} and transitions {t1, t2, t3, t4, t5}. The pre-set of t2 is {s1, s6}. The post-set of {s3, s4, s5} is {t3, t4, t5}. The sequence t2 s2 t1 s1 t2 s3 is a path. This net is strongly connected. The net obtained by removing t1 and its adjacent arcs is connected but not strongly connected. The net obtained by removing t2 and its adjacent arcs is not connected.

The following proposition follows immediately from the definitions.

Proposition 1. *A weakly connected net is strongly connected if and only if for each directed arc* (x, y) *there is a directed path leading from* y *to* x. □

3.2 Markings and Behavior

The states of a Petri net are defined by its markings, to be defined next. Elementary state changes are caused by the occurrences of transitions. The *occurrence rule* distinguishes the *enabling condition* and the *marking transformation.*

A *marking* of a net N is a mapping $m: S_N \to I\!N$ where $I\!N = \{0, 1, 2, \dots\}$. A place s is *marked* by a marking m if $m(s) > 0$. The *null marking* is the marking which maps every place to 0.

A transition t is *enabled* by a marking m if m marks all places in ${}^\bullet t$. In this case t can *occur.* Its occurrence transforms m into the marking m', defined for each place s by

$$m'(s) = \begin{cases} m(s) - 1 & \text{if } s \in {}^\bullet t - t^\bullet, \\ m(s) + 1 & \text{if } s \in t^\bullet - {}^\bullet t, \\ m(s) & \text{otherwise.} \end{cases}$$

Notice that a place in ${}^\bullet t \cap t^\bullet$ is marked whenever t is enabled but does not change its token count by occurrence of t. Sometimes we will use the following equivalent equation for the marking m':

$$m'(s) = m(s) - |F_N \cap \{(s,t)\}| + |F_N \cap \{(t,s)\}|.$$

A marking is called *dead* if it enables no transition of N.

Consider the net shown in Figure 9 with the depicted marking, mapping `s1`, `s2`, and `s3` to 1 and all other places to 0. Transition `t1` is enabled, and the marking reached after its occurrence maps `s1` to 2 and `s2` to 0. Transition `t3` is enabled, too. The null marking is an example for a dead marking.

3.3 Occurrence sequences

Let m be a marking of a net. A finite sequence $t_1 \ldots t_k$ of transitions is called a *finite occurrence sequence, enabled at* m, if there are markings $m_1, \ldots, m_k$ such that

$$m \xrightarrow{t_1} m_1 \xrightarrow{t_2} \cdots \xrightarrow{t_k} m_k.$$

We write $m \xrightarrow{\sigma} m_k$ in this case, where $\sigma = t_1 \ldots t_k$. The empty sequence ε is enabled at any marking m and satisfies $m \xrightarrow{\varepsilon} m$. A marking m' is *reachable* from a marking m if there is a finite occurrence sequence $m \xrightarrow{\sigma} m'$.

An infinite sequence $t_1\, t_2\, t_3 \ldots$ is called *infinite occurrence sequence, enabled at* m, if there are markings $m_1, m_2, \ldots$ such that

$$m \xrightarrow{t_1} m_1 \xrightarrow{t_2} m_2 \xrightarrow{t_3} \cdots.$$

The following propositions follow immediately from the definitions.

Proposition 2. *If $m \xrightarrow{\sigma} m'$ is a finite occurrence sequence and m' enables a (finite or infinite) occurrence sequence σ', then the sequence $\sigma\, \sigma'$ is enabled at m, too.* □

Proposition 3. *An infinite sequence σ of transitions is enabled at a marking m if and only if every finite prefix of σ is enabled at m.* □

Proposition 4. *If $m \xrightarrow{\sigma} m'$ and $l \xrightarrow{\sigma} l'$, then $m'(s) - m(s) = l'(s) - l(s)$ for each place s.* □

Using the previous two propositions and the occurrence rule, it is easy to prove the following result.

Proposition 5. *If m and l are markings satisfying $m(s) \geq l(s)$ for each place s then every occurrence sequence enabled at l is also enabled at m.* □

Occurrence sequences do not provide full information about the causal relationships between transitions. Assume an occurrence sequence $v\, u$, enabled at a marking m. Transition u might be enabled only after the occurrence of v, or v and u might be *concurrently* enabled at m. In the latter case, $u\, v$ is an occurrence sequence, too. The following lemma identifies a sufficient condition for exchanging concurrent transitions.

Lemma 6. *Assume that u and v are transitions of a net satisfying $v^\bullet \cap {}^\bullet u = \emptyset$. If $m_1 \xrightarrow{v\,u} m_3$ then $m_1 \xrightarrow{u\,v} m_3$.*

Proof. Let $m_1 \xrightarrow{v} m_2$. We have $m_1(s) \geq m_2(s)$ for $s \in {}^\bullet u$ because $v^\bullet \cap {}^\bullet u = \emptyset$. Moreover, $m_2(s) \geq 1$ for $s \in {}^\bullet u$ because m_2 enables u. Therefore, m_1 enables u.

We have $m_1(s) \geq 1$ for $s \in {}^\bullet v$ because m_1 enables v. Moreover, $m_1(s) \geq 2$ for $s \in {}^\bullet u \cap {}^\bullet v$ because m_1 enables $v\,u$ and $v^\bullet \cap {}^\bullet u = \emptyset$. Therefore, m_1 enables $u\,v$.

By Proposition 4, the sequences $u\,v$ and $v\,u$ cause the same marking change. Therefore, the occurrence of $u\,v$ at m_1 leads to the marking m_3, too. □

3.4 Marked Nets and their Properties

A *marked net* is a net equipped with a marking, called *initial marking*. Throughout this paper we always denote the initial marking of a net by m_0. The initial marking of a marked net is graphically depicted by tokens in the places, as shown in Figure 9.

An *occurrence sequence of a marked net* is an occurrence sequence enabled by the initial marking. The *reachable markings* of a marked net are the markings reachable from the initial marking (including the initial marking itself).

Let $b \in \mathbb{N}$. A place s of a marked net is called *bounded* with bound b (*b-bounded*) if $m(s) \leq b$ for each reachable marking m. It is *bounded* if it is b-bounded for some b. A marked net is called *b-bounded* (*bounded*) if all its places are b-bounded (bounded, respectively).[5]

Proposition 7. *A finite marked net is bounded if and only if there is a bound b such that all its places are b-bounded.* □

Notice that this property fails for infinite nets because there might be infinitely many different bounds.

A transition t is called *dead* at a marking m if it is not enabled at any marking reachable from m. A marked net is said to be *live* if there is no dead transition at any reachable marking. It is *deadlock-free* if no reachable marking is dead.

Proposition 8. *Each live marked net with at least one transition is deadlock-free.* □

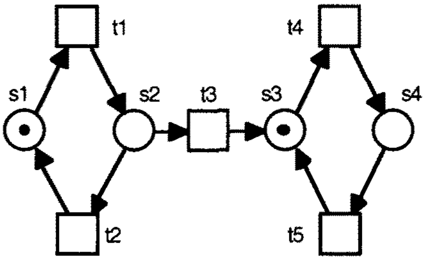
Notice that each marked net without transitions is live but not deadlock-free.

Often it is required that every place of a marked net eventually can be marked.

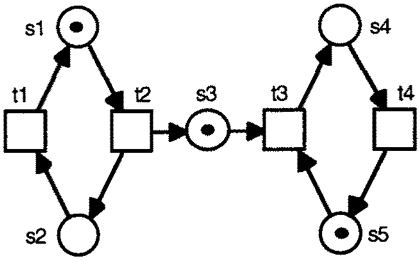
Proposition 9. *If no transition of a marked net is dead then each non-isolated place is marked at some reachable marking.* □

A marked net is called *reversible* if its initial marking can be reached from any other reachable marking.

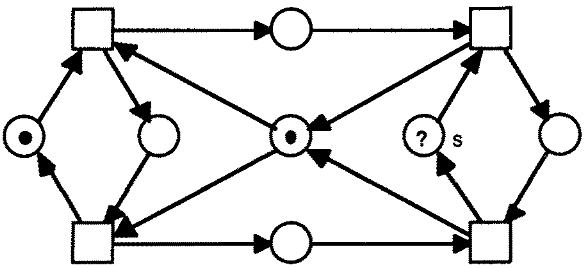
[5] 1-bounded places are sometimes called 1-safe or just safe places. Accordingly, marked nets with only 1-bounded places are sometimes called 1-safe or safe marked nets.



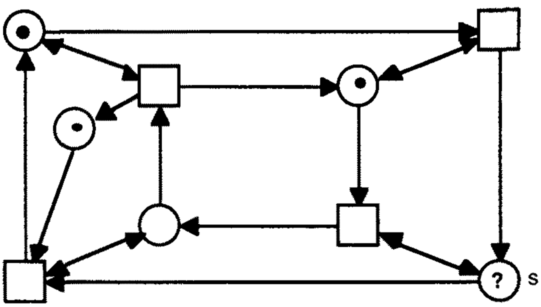
N1: bounded, not live



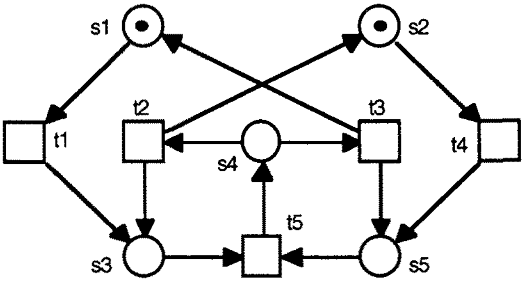
N2: live, not bounded



N3: bounded, live only if $m_0(s) = 1$



N4: live, bounded only if $m_0(s) = 0$



N5: live, 1-bounded, not reversible

Fig. 10. Examples for liveness, boundedness and reversibility of marked nets

Proposition 10. *A reversible marked net is deadlock-free if and only if the initial marking is not dead.* □

The marked net shown in Figure 9 is live, 2-bounded and reversible. Figure 10 presents further examples for the above notions. These examples also show that neither liveness of bounded marked nets nor boundedness of live marked nets is monotonous in the sense that the property is preserved when tokens are added to places.

3.5 The Strongly-Connectedness-Theorem

Live and bounded marked nets enjoy many interesting properties. One important example is the Strongly-Connectedness-Theorem which states that such a net is strongly connected, provided it is weakly connected. This result implies that each of the first two nets of Figure 10 either is not live or not bounded, for any initial marking:

Theorem 11. *Every weakly connected live and bounded marked net is strongly connected.*

Proof:[6] Assume a weakly connected live and bounded marked net. By Proposition 1 it suffices to prove for each arc (x, y) that there is a directed path leading from y to x.

Case 1: x is a place and y is a transition. Let V be the set of all transitions t for which there is a directed path from y to t. Let U be the set of all other transitions of the net. Then $u^\bullet \cap {}^\bullet v = \emptyset$ for every u in U and v in V.

Let b be the bound of place y. Liveness implies that the initial marking enables a finite occurrence sequence σ with $b+1$ occurrences of transition y. By Lemma 6, transitions of U and of V can repeatedly be swapped, resulting in an occurrence sequence $\sigma_1\,\sigma_2$, such that σ_1 contains all occurrences of transitions of U in σ and σ_2 contains all occurrences of transitions of V in σ.

Transition y is in the set V by definition of V. So y occurs $b+1$ times in σ_2. Since x is b-bounded and $y \in x^\bullet$, some transition of ${}^\bullet x$ occurs in σ_2. Since σ_2 contains only transitions of V, there is a transition in ${}^\bullet x \cap V$. By definition of V, there is a directed path from y to that transition. This path can be extended by x.

Case 2: x is a transition and y is a place. Let U be the set of all transitions t for which there is a directed path from t to x. Let V be the set of all other transitions of the net. Then $u^\bullet \cap {}^\bullet v = \emptyset$ for every u in U and v in V.

Let b be the bound of place y. Liveness implies that the initial marking enables a finite occurrence sequence σ with $b+1$ occurrences of transition x. We continue as in Case 1: By Lemma 6, transitions of U and of V can repeatedly be swapped, resulting in an occurrence sequence $\sigma_1\,\sigma_2$, such that σ_1 contains all

[6] This proof employs occurrence sequences and Lemma 6. For a more elegant proof based on partially ordered processes see Section 6.

occurrences of transitions of U in σ and σ_2 contains all occurrences of transitions of V in σ.

Transition x is in the set U by definition of U. So x occurs $b+1$ times in σ_1. Since y is b-bounded and $x \in {}^\bullet y$, some transition of $y^\bullet$ occurs in σ_1. Since σ_1 contains only transitions of U, there exists a transition in $y^\bullet \cap U$. By definition of U, there is a directed path from that transition to x. Concatenation of y and this path yields a directed path from y to x. □

4 Marking Graphs and Coverability Graphs

4.1 Marking Graphs

The sequential behavior of a marked net is given by its set of occurrence sequences. Each sequence of transition occurrences

$$m_0 \xrightarrow{t_1} m_1 \xrightarrow{t_2} m_2 \xrightarrow{t_3} \cdots,$$

can be considered a linear directed, acyclic graph

$$\overset{m_0}{\bullet} \xrightarrow{t_1} \overset{m_1}{\bullet} \xrightarrow{t_2} \overset{m_2}{\bullet} \xrightarrow{t_3} \cdots$$

with vertices labeled by reachable markings and edges labeled by transitions. In this way, the set of all occurrence sequences yield an acyclic graph consisting of many disconnected components: one component for each occurrence sequence. Equally labeled verices may be identified, yielding a connected graph, which is not acyclic in general. Its vertices correspond bijectively to the reachable markings of the marked net. Each transition enabled at a marking m yields one outgoing edge of the corresponding vertex. A distinguished vertex represents the initial marking. Each directed path starting at this vertex uniquely represents an occurrence sequence. Vice versa, each occurrence sequence is uniquely represented by a path.

Figures 11 shows the marking graphs of the marked nets **N1**, **N2** and **N5** given in Figure 10. In these graphs, each marking m is denoted by a vector such that the i-th entry of a vector is $m(\texttt{si})$. Initial markings are depicted by arrows without source.
Formal definition of marking graphs requires some notions about graphs:

An *arc-labeled directed graph* is given by

- a set V of *vertices* and
- a set of *labeled edges* (v, l, v') (representing source vertex, label, and target vertex, respectively) where $v, v' \in V$ and l is a label of some given set L.

If (v, l, v') is a labeled edge then v' is an *immediate successor* of v.

A *path* of an arc-labeled graph is a (finite or infinite) sequence of labeled edges

$$(v, l_1, v_1), (v_1, l_2, v_2), (v_2, l_3, v_3), \ldots.$$

All vertices in $\{v, v_1, v_2, v_3, \ldots\}$ are called *successors* of v.

The *marking graph* of a marked net is an arc-labeled directed graph with a distinguished initial vertex and edges labeled by transitions:

- the *vertices* are the reachable markings,
- the distinguished *initial vertex* is the initial marking,
- labeled edges are all triples (m, t, m') such that m and m' are reachable markings satisfying $m \xrightarrow{t} m'$.

4.2 Marking Graphs and Properties of Marked Nets

Proposition 12. *Given a marked net with initial marking m_0,*

$$m_0 \xrightarrow{t_1} m_1 \xrightarrow{t_2} m_2 \xrightarrow{t_3} \cdots$$

is a (finite or infinite) sequence of subsequent transition occurrences if and only if there is a path

$$(m_0, t_1, m_1), (m_1, t_2, m_2), (m_2, t_3, m_3), \ldots$$

of the marking graph. □

Behavioral properties of marked nets are closely related to structural properties of their marking graphs, as stated in the following results.

Lemma 13. *A finite marked net is bounded if and only if its marking graph has finitely many vertices.*

Proof. By definition, the set of vertices of the marking graph is finite if and only if the set of reachable markings is finite.
($\Longleftarrow$) If a marked net with initial marking m_0 has k reachable markings then any reachable marking can be reached by at most $k-1$ transition occurrences. Since a transition occurrence increases the number of tokens on a place by at most one, each place s is bounded by $m_0(s) + k - 1$.
($\Longrightarrow$) Assume that the net has n places. By Proposition 7 there is a common bound b for all places. Hence each place s has no more than $b+1$ different states, given by $m(s) = 0, \ldots, m(s) = b$. So there are no more than $(b+1)^n$ reachable markings. □

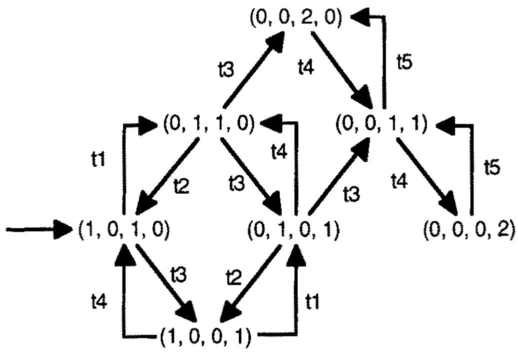
Lemma 14. *A marked net is deadlock-free if and only if its marking graph has no vertex without outgoing edge.*

Proof. Follows immediately from Proposition 12. □

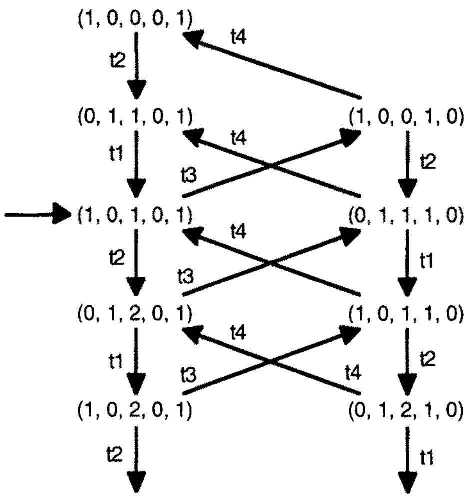
Lemma 15. *A finite marked net is live if and only if, for each vertex m of its marking graph, there exists a path*

$$(m, t_1, m_1), (m_1, t_2, m_2), \ldots, (m_{k-1}, t_k, m_k)$$

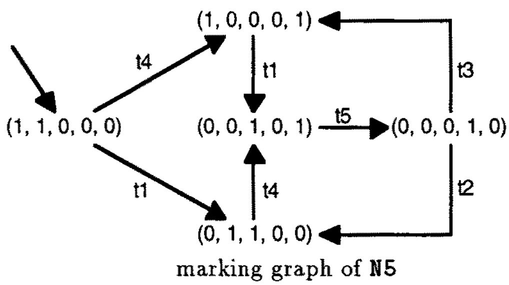
such that the sequence $t_1\, t_2 \ldots t_k$ contains all transitions of the net.



marking graph of **N1**



initial part of the marking graph of **N2**



marking graph of **N5**

Fig. 11. Marking graphs of marked nets given in Figure 10

Proof.
($\Longleftarrow$) Each reachable marking m is represented by a vertex of the marking graph. By Proposition 12, the path $(m, t_1, m_1), (m_1, t_2, m_2), \ldots, (m_{k-1}, t_k, m_k)$ corresponds to an occurrence sequence $m \xrightarrow{t_1 \ldots t_k} m_k$. So each transition is enabled by some marking reachable from m.
($\Longrightarrow$) Let m be a reachable marking. Let $t_1, \ldots, t_m$ be a list of all transitions of the net. Liveness implies that some marking reachable from m enables t_1.
Let $m \xrightarrow{\sigma_1} m' \xrightarrow{t_1} m_1$.
Again by liveness, we find an occurrence sequence $m_1 \xrightarrow{\sigma_2} m_1' \xrightarrow{t_2} m_2$.
Repeated application of this construction eventually leads to an occurrence sequence that contains every transition at least once. The sequence of edge labels of the corresponding path of the marking graph contains all transitions. □

Lemma 16. *A marked net is reversible if and only if its marking graph is strongly connected.* [7]

Proof.
($\Leftarrow$) Since the marking graph is strongly connected, there are directed paths from any vertex to the initial vertex. Hence the initial marking can be reached from any reachable marking.
($\Rightarrow$) Let m_1 and m_2 be arbitrary reachable markings. We show that the marking graph has a directed path from m_1 to m_2. By reversibility of the marked net, there is an occurrence sequence from m_1 to the initial marking m_0. Since m_2 is reachable, there is an occurrence sequence from m_0 to m_2. The concatenation of these sequences leads from m_1 to m_2. The corresponding path of the marking graph leads from the vertex m_1 to the vertex m_2 □

The marking graphs of Figure 11 provides examples. The first one represents the behavior of a bounded deadlock-free marked net which is not live. The second one stems from of an unbounded live marked net. The marked net of the third marking graph is live and bounded but not reversible.

4.3 ω-Markings and ω-Occurrence Sequences

By Lemma 13, unbounded finite marked nets have infinitely many reachable markings and thus infinite marking graphs. Therefore, analysis techniques based on the marking graph are not applicable to them. A finite graph can be constructed instead, providing important information about the behavior of the marked net. In particular, this graph can be used to decide boundedness of places and deadness of transitions. The graph is actually a tree, called *coverability tree*.

The core idea of the following sections is the concept of *ω-markings* which generalize conventional markings.

Formally, an *ω-marking* of a net N is a mapping $\overline{m}: S_N \rightarrow \mathbb{N} \cup \{\omega\}$ where $\omega \notin \mathbb{N}$. Clearly, every (conventional) marking can be viewed as a particular ω-marking without ω-entries.

[7] For that reason, reversible marked nets are sometimes called *cyclic*.

ω-markings are interpreted as follows: If a marking m' is reachable from a marking m and satisfies $m'(s) \geq m(s)$ for each place s, the occurrence sequence leading from m to m' can be iterated arbitrarily often (Proposition 5). If moreover $m'(s_0) > m(s_0)$ for some place s_0 then the number of tokens on s_0 increases with each iteration of the occurrence sequence. This increasing sequence of markings is now replaced by one ω-marking $\overline{m'}$ with $\overline{m'}(s_0) = \omega$, denoting that, for each $b \in I\!N$, there is a reachable marking that coincides with m' for all places except s_0 and assigns at least b tokens to s_0. More generally, several places may map to ω, representing simultaneous growth of the token count on these places.

ω-markings are constructed by help of ω-occurrence sequences. The following definition of ω-occurrence sequences demonstrates that the occurrence rule can be applied for ω-markings in a similar way as for usual markings. An ω-entry of a place is not affected by any transition occurrence. Moreover, no ω-marking, except occasionally the last one, should occur twice in an ω-occurrence sequence. We will first define finite and infinite ω-occurrence sequences and then prove that every ω-occurrence sequence of a finite marked net is finite.

A (finite or infinite) sequence of transitions $t_1\, t_2\, t_3 \ldots$ is an *ω-occurrence sequence* of a marked net with initial marking m_0 if there exist ω-markings $\overline{m}_0, \overline{m}_1, \overline{m}_2, \ldots$ such that m_0 and $\overline{m}_0$ coincide for all places and, for each index i occurring in the sequence $t_1\, t_2\, t_3 \ldots$ the following conditions hold:

(1) For each place s in ${}^\bullet t_i$, either $\overline{m}_{i-1}(s) > 0$ or $\overline{m}_{i-1}(s) = \omega$ (the enabling condition).

(2) For each place s satisfying $\overline{m}_i(s) \neq \omega$,

$$\overline{m}_i(s) = \overline{m}_{i-1}(s) - |F_N \cap \{(s, t_i)\}| + |F_N \cap \{(t_i, s)\}|$$

(the conventional marking transformation).

(3) A place s satisfies $\overline{m}_i(s) = \omega$ if and only if
   - either $\overline{m}_{i-1}(s) = \omega$
     (places marked by ω remain marked by ω),
   - or $\overline{m}_{i-1}(s) \neq \omega$ and there exists an index j, $j < i$, such that $\overline{m}_j(s) \neq \omega$ and $\overline{m}_j(s) < \overline{m}_{i-1}(s) - |F_N \cap \{(s, t_i)\}| + |F_N \cap \{(t_i, s)\}|$ and $\overline{m}_j(s') \leq \overline{m}_{i-1}(s') - |F_N \cap \{(s', t_i)\}| + |F_N \cap \{(t_i, s')\}|$ for each place s' satisfying $\overline{m}_j(s') \neq \omega$ and $\overline{m}_{i-1}(s') \neq \omega$
     (places with increasing token count are marked by ω).

(4) If $i > 1$ then $\overline{m}_{i-1} \notin \{\overline{m}_0, \ldots, \overline{m}_{i-2}\}$
(after reaching an ω-marking the second time, the sequence stops).

We call an ω-marking $\overline{m}$ *reachable* in a marked net if some ω-occurrence sequence leads to $\overline{m}$.

Figure 12 shows an unbounded marked net. A part of its marking graph is given in Figure 13 (all markings with less than 5 tokens are shown). The ω-occurrence sequences of this marked net are:

t1 t1, **t1 t2 t3 t3**, **t1 t2 t3 t4 t1**, **t1 t2 t4**, **t2 t4**,

as well as all prefixes of these sequences.

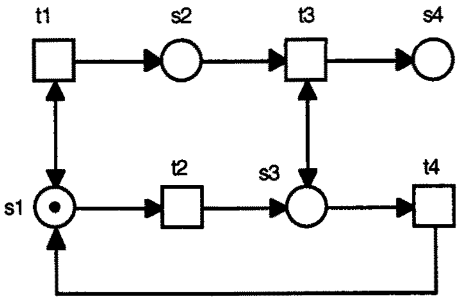


Fig. 12. An unbounded marked Petri net

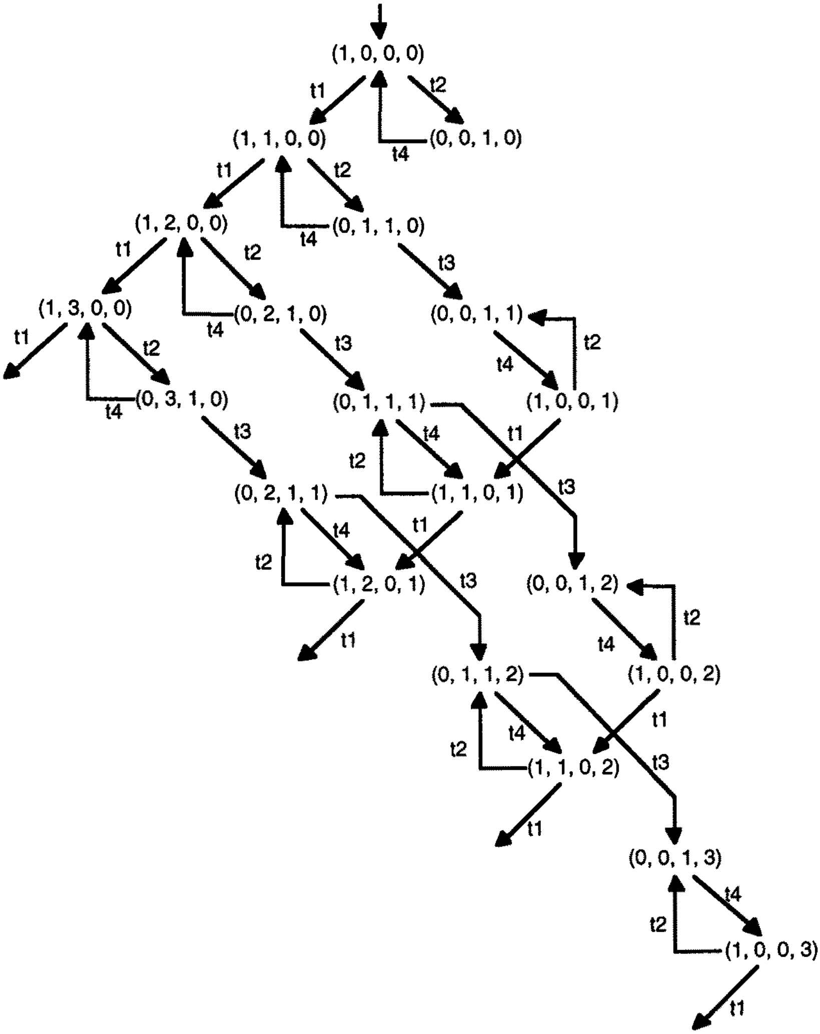


Fig. 13. A part of the marking graph of the marked net of Figure 12

Note that `t1 t2 t3 t3` is an ω-occurrence sequence but not an occurrence sequence. Conversely, `t1 t1 t1` is an example for an occurrence sequence which is not an ω-occurrence sequence, although no marking is reached twice during this sequence.

4.4 Finiteness of ω-Occurrence Sequences

The following theorem states that every ω-occurrence sequence of a finite marked net is finite. For technical convenience, we assume a linear order $\leq$ on the set $\mathbb{N} \cup \{\omega\}$, given by $x \leq y$ if and only if $y = \omega$ or $x \leq y$ in $\mathbb{N}$. The following lemma is known as *Dickson's Lemma.*

Lemma 17. *Let S be a finite set and let $\varphi_1\, \varphi_2\, \varphi_3 \ldots$ be an infinite sequence of mappings from S to $\mathbb{N} \cup \{\omega\}$. There exists an infinite sequence of indices $i_1\, i_2\, i_3\, \ldots$ which is strongly monotonic (i.e., $i_1 < i_2 < i_3 < \cdots$) such that, for each s in S,*

$$\varphi_{i_1}(s) \leq \varphi_{i_2}(s) \leq \varphi_{i_3}(s) \leq \cdots.$$

Proof. We prove the following stronger proposition: For each subset S' of S, there exists an infinite strongly monotonic sequence of indices $i_1, i_2, i_3, \ldots$ such that, for each s in S', $\varphi_{i_1}(s) \leq \varphi_{i_2}(s) \leq \varphi_{i_3}(s) \leq \cdots$. We proceed by induction on the number of elements in S'.
Base. If $S' = \emptyset$ then nothing has to be shown.
Step. Assume $S' \neq \emptyset$ and let $s \in S'$. By the induction hypothesis, there exists an infinite strongly monotonic sequence $i_1, i_2, i_3, \ldots$ such that, for each s' in $S' \backslash \{s\}$,

$$\varphi_{i_1}(s') \leq \varphi_{i_2}(s') \leq \varphi_{i_3}(s') \leq \cdots.$$

Now we restrict the sequence $i_1, i_2, i_3, \ldots$ to indices i_k satisfying

$$\varphi_{i_k}(s) \leq \varphi_{i_{k+1}}(s),\;\; \varphi_{i_k}(s) \leq \varphi_{i_{k+2}}(s),\;\; \varphi_{i_k}(s) \leq \varphi_{i_{k+3}}(s) \ldots$$

Clearly, the obtained sequence $i_{k_1}, i_{k_2}, i_{k_3}, \ldots$ satisfies the required property

$$\varphi_{i_{k_1}}(s) \leq \varphi_{i_{k_2}}(s) \leq \varphi_{i_{k_3}}(s) \leq \cdots$$

for each place s in S'. This sequence is infinite because, for each index i_k, every index i_l in $\{i_{k+1}, i_{k+2}, i_{k+3} \ldots\}$ satisfying

$$\varphi_{i_l}(s) \leq \varphi_{i_{k+1}}(s), \varphi_{i_{k+2}}(s), \varphi_{i_{k+3}}(s) \ldots$$

belongs to the sequence, too. Such an index i_l always exists because every nonempty subset of $\mathbb{N} \cup \{\omega\}$ has a minimal element. □

Theorem 18. *Every ω-occurrence sequence of a finite marked net is finite.*

Proof. By contraposition, assume a finite marked net that has an infinite ω-occurrence sequence $t_1\, t_2\, t_3 \ldots$,

$$\overline{m}_1 \xrightarrow{t_1} \overline{m}_2 \xrightarrow{t_2} \overline{m}_3 \xrightarrow{t_3} \cdots.$$

By Dickson's Lemma (Lemma 17), there exists an infinite strongly monotonic sequence of indices $i_1, i_2, i_3 \ldots$ such that, for each place s,

$$\overline{m}_{i_1}(s) \leq \overline{m}_{i_2}(s) \leq \overline{m}_{i_3}(s) \leq \cdots.$$

Let i and j be two subsequent indices of the sequence $i_1, i_2, i_3 \ldots$. By the definition of ω-occurrence sequences (4) no ω-marking appears twice in an infinite ω-occurrence sequence. Hence $\overline{m}_i(s) \neq \overline{m}_j(s)$ for at least one place s. By the definition of ω-occurrence sequences (3), $\overline{m}_i(s) \neq \omega$ and $\overline{m}_j(s) = \omega$. Again by (3), no place s satisfies $\overline{m}_i(s) = \omega$ and $\overline{m}_j(s) \neq \omega$. Hence $\overline{m}_j$ has more places with ω-entries than $\overline{m}_i$. Therefore, the set of places with ω-entries increases infinitely, contradicting the finiteness of the set of all places of the net. □

4.5 Coverability Trees

Next we show that a finite marked net has only finitely many ω-occurrence sequences and hence, together with the previous result, only finitely many reachable ω-markings. To this end, we introduce coverability trees.

As shown before for occurrence sequences, every ω-occurrence sequence can be conceived as a linear directed, acyclic graph with vertices labeled by ω-markings and edges labeled by transitions. Starting with a single disconnected graph representing all ω-occurrence sequences of a marked net, we identify common prefixes of ω-occurrence sequences. More precisely, we identify two vertices of the graph if and only if they are reached by the same ω-occurrence sequence. This construction yields the coverability tree of the marked net.

Formally, the *coverability tree* of a marked net is defined as a directed graph with a distinguished initial vertex and edges labeled by transitions:

- the vertices are the finite ω-occurrence sequences,
- a distinguished initial vertex is given by the empty sequence ε (which by definition is an ω-occurrence sequence),
- labeled edges are all triples $(\sigma,\ t,\ \sigma\, t)$ such that σ as well as $\sigma\, t$ are ω-occurrence sequences.

It is obvious that the coverability tree is a tree with the empty sequence ε being the root. In the graphical representation, any vertex σ is labeled by the ω-marking reached by σ.

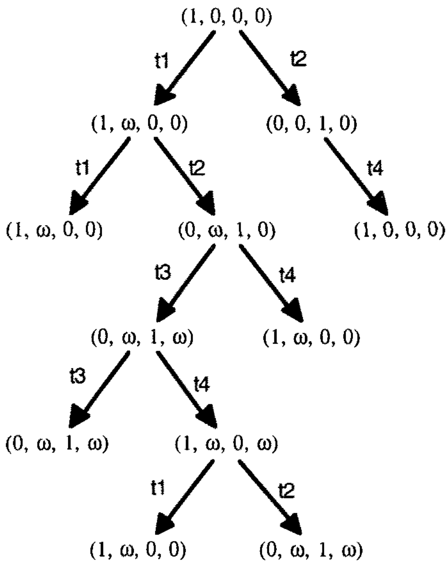


Fig. 14. The coverability tree of the marked net of Figure 12

Figure 14 shows the coverability tree of the marked net of Figure 12. Each vertex is the ω-occurrence sequence given by the arc labels of the path leading from the root to that vertex. The vertices are annotated by the respective reached ω-markings (in vector notation).

Theorem 19. *The coverability tree of a finite marked net is finite.*[8]

Proof. By contraposition, assume a finite marked net with an infinite coverability tree. Each vertex σ of the coverability tree has only finitely many immediate successors, one for each transition enabled by the ω-marking reached by σ. Hence every vertex σ with infinitely many successors has at least one immediate successor which also has infinitely many successors. By assumption, the initial vertex ε has infinitely many successors. Hence, starting with ε, we can construct an infinite directed path of the tree. The concatenation of the labels of the edges of this path yields an infinite ω-occurrence sequence — contradicting Theorem 18. □

Corollary 20. *A finite marked net has finitely many reachable ω-markings.* □

4.6 Reachable Markings and Reachable ω-Markings

The following two theorems formulate relations between reachable markings and reachable ω-markings of a marked net.

[8] This result is based on the well-known lemma of König

Theorem 21. *Let $\overline{m}$ be a reachable ω-marking of a finite marked net. For each b in $\mathbb{N}$, there is a reachable marking m such that every place s satisfies:*

- *if $\overline{m}(s) \neq \omega$ then $m(s) = \overline{m}(s)$,*
- *if $\overline{m}(s) = \omega$ then $m(s) \geq b$.*

Proof. Let N be a net with initial marking m_0. Let

$$\overline{m}_0 \xrightarrow{t_1} \overline{m}_1 \xrightarrow{t_2} \cdots \xrightarrow{t_k} \overline{m}_k$$

be an ω-occurrence sequence. By definition, $m_0(s) = \overline{m}_0(s)$ for each place s.

Let S_ω be the set of places satisfying $\overline{m}_k(s) = \omega$. Let $b \in \mathbb{N}$; we prove that there exists a reachable marking m which coincides with $\overline{m}_k$ on all places not in S_ω and assigns at least b tokens to each place in S_ω. To this end, we construct an occurrence sequence

$$m_0 \xrightarrow{t_1} m_1 \xrightarrow{\sigma_1} m_1' \xrightarrow{t_2} m_2 \xrightarrow{\sigma_2} m_2' \xrightarrow{t_3} \cdots \xrightarrow{\sigma_{k-1}} m_{k-1}' \xrightarrow{t_k} m_k \xrightarrow{\sigma_k} m_k'$$

such that m_k' satisfies the above property.

Let s be an arbitrary place in S_ω. By definition of ω-occurrence sequences, there is a unique index $i(s)$ such that

$$\overline{m}_0(s), \overline{m}_1(s), \ldots, \overline{m}_{i(s)-1}(s) \neq \omega,$$

$$\overline{m}_{i(s)}(s), \overline{m}_{i(s)+1}(s), \ldots, \overline{m}_k(s) = \omega$$

and an index $j(s)$, $j(s) < i(s)$, such that the transition sequence $\tau(s)$,

$$\tau(s) = t_{j(s)+1}\, t_{j(s)+2} \cdots t_{i(s)},$$

increases the token count of s and does not change the token count of any place s' satisfying $\overline{m}_{i(s)}(s') \neq \omega$. However, the sequence $\tau(s)$ might decrease the token count of places s' satisfying $\overline{m}_{i(s)-1}(s') = \omega$. It decreases the token count of a place s by at most $|\tau(s)|$, the length of $\tau(s)$. Any marking that coincides with $\overline{m}_{i(s)}$ on each place s' satisfying $\overline{m}_{i(s)}(s') \neq \omega$ and associates at least $|\tau(s)|$ tokens to each other place, enables $\tau(s)$.
For $n = 1, \ldots, k$, let $S_n = \{s \in S_\omega | i(s) = n\}$ and assume $S_n = \{s_n^1, \ldots, s_n^{|S_n|}\}$. In the sequel, $[\tau]^m$ stands for the sequence $\underbrace{\tau\, \tau \ldots \tau}_{m \text{ times}}$.

Define

$$\sigma_k = \left[\tau(s_k^1)\, \tau(s_k^2)\, \ldots\, \tau(s_k^{|S_k|})\right]^b$$

to ensure that each place s in S_k satisfies $m_k'(s) \geq b$. Since the token count of places in S_{k-1} can be decreased by t_k and by σ_k, define

$$\sigma_{k-1} = \left[\tau(s_{k-1}^1)\, \tau(s_{k-1}^2)\, \ldots\, \tau(s_{k-1}^{|S_k|})\right]^{(|\sigma_k|+1+b)}$$

to ensure that $t_k\, \sigma_k$ is enabled after the occurrence of σ_{k-1} and moreover leaves at least b tokens on each place in S_{k-1}.

In general, for $i = k, k-1, \ldots, 1$ the sequence σ_i is defined by

$$\sigma_i = \left[\tau(s_i^1)\,\tau(s_i^2)\,\ldots\,\tau(s_i^{|S_i|})\right]^{(|\sigma_{i+1}|+\cdots+|\sigma_k|+(k-i)+b)} .$$

Notice that some of the sequences σ_i might be empty. Since the marking $\overline{m}_0$ has no ω-entries, the sequence σ_1 does not decrease the token count of any place and is hence enabled at m_1.

Since the token count of a place not in S_ω is not changed by any occurrence sequence τ_i, it is not changed by any σ_i. Hence m_k' and $\overline{m}_k$ coincide on all these places. □

Theorem 22. *Let m be a reachable marking of a finite marked net. There exists a reachable ω-marking $\overline{m}$ such that $m(s) = \overline{m}(s)$ for each place s satisfying $\overline{m}(s) \neq \omega$.*

Proof. A sequence of transitions is called *generalized ω-occurrence sequence* of a marked net if it satisfies conditions (1) to (3) of the definition of ω-occurrence sequences, i.e., generalized ω-occurrence sequences are very much like ω-occurrence sequences except they do not necessarily stop after reaching an ω-marking twice.

We claim that, for each generalized ω-occurrence sequence σ leading to an ω-marking $\overline{m}$, there also exists an ω-occurrence sequence leading to $\overline{m}$. The claim is proven by induction on the length k of σ.
Base: $k = 0$. The empty sequence is an ω-occurrence sequence.
Step: $k > 0$. If σ is an ω-occurrence sequence then there is nothing to be shown. So assume that this is not the case. Let $\sigma = t_1\,t_2\,\ldots\,t_k$ and

$$\overline{m}_0 \xrightarrow{t_1} \overline{m}_1 \xrightarrow{t_2} \cdots \xrightarrow{t_k} \overline{m}_k .$$

Since σ is not an ω-occurrence sequence, there is a smallest index i $(1 \leq i \leq k)$ such that $\overline{m}_i$ coincides with some $\overline{m}_j$ $(0 \leq j < i)$. The sequence $t_1 \ldots t_j t_{i+1} \ldots t_k$ is a generalized ω-occurrence sequence, too. The ω-marking reached by this sequence is $\overline{m}_k$, too. Since the length of this sequence is smaller than k, the induction hypothesis can be applied, which finishes proof of the claim.

Let $m_0 \xrightarrow{t_1} m_1 \xrightarrow{t_2} \cdots \xrightarrow{t_k} m_k$ be an occurrence sequence such that $m_k = m$. Then, obviously $t_1\,t_2 \ldots t_k$ is a generalized ω-occurrence sequence. This generalized ω-occurrence sequence leads to an ω-marking $\overline{m}$ satisfying the required property. By the above claim, some ω-occurrence sequence leads to $\overline{m}$, too. □

In turn, the previous result implies a sufficient condition for non-reachability of a marking.

4.7 Analysis Through Coverability Trees

By construction of the coverability tree of a marked net, we obtain a finite representation of all reachable ω-markings. The following theorems show that by inspection of this finite set of ω-markings, boundedness of places and deadness of transitions can be decided.

Theorem 23. *A place s of a marked net is not bounded if and only if some reachable ω-marking $\overline{m}$ satisfies $\overline{m}(s) = \omega$ (i.e., some vertex of the coverability tree represents the ω-marking $\overline{m}$).*

Proof.
($\Longleftarrow$) follows immediately from Theorem 21.
($\Longrightarrow$) Since there are only finitely many reachable ω-markings by Theorem 19 there is a number $b \in \mathbb{N}$ such that each reachable ω-marking $\overline{m}$ satisfies either $\overline{m}(s) = \omega$ or $\overline{m}(s) < b$. Since s is not bounded, some reachable marking m satisfies $m(s) \geq b$. Since $m(s)$ does not coincide with $\overline{m}(s)$ for any reachable ω-marking $\overline{m}(s)$, there exists some reachable ω-marking $\overline{m}$ satisfying $\overline{m}(s) = \omega$ by Theorem 22. □

Together with Theorem 22, this result implies the following corollary.

Corollary 24. *A place s of a marked net is b-bounded if and only if each reachable ω-marking $\overline{m}$ satisfies $\overline{m}(s) \neq \omega$ and $\overline{m}(s) \leq b$.*

Theorem 25. *A transition t of a marked net is dead if and only if t does not occur in any ω-occurrence sequence (i.e., some arc of the coverability tree is labeled by t).*

Proof.
($\Longleftarrow$) Assume some reachable marking m enables t. By Theorem 22, a corresponding reachable ω-marking $\overline{m}$ satisfies $\overline{m}(s) \neq 0$ for each place s in ${}^{\bullet}t$. Hence, this ω-marking enables t, too.
($\Longrightarrow$) Assume some reachable ω-marking $\overline{m}$ enables t. By Theorem 21, there is a corresponding reachable marking m that marks all places satisfying $\overline{m} = \omega$ at least once. This marking m enables t, too. □

4.8 Coverability Trees and Marking Graphs

Theorem 23 immediately implies the following corollary.

Corollary 26. *A marked net is bounded if and only if no reachable ω-marking contains an ω-entry.*

In other words: reachable markings and reachable ω-markings coincide for bounded marked nets.

The *coverability graph* of a marked net is defined as an arc-labeled directed graph with a distinguished initial vertex and edges labeled by transitions:

- the *vertices* are the reachable ω-markings,
- the distinguished *initial vertex* is given by the ω-marking that coincides with the initial marking for each place,
- labeled edges are given by all triples $(\overline{m}, t, \overline{m}')$ such that $\overline{m}$ and $\overline{m}'$ are reachable ω-markings satisfying $\overline{m} \xrightarrow{t} \overline{m}'$.

The coverability graph is easily constructed from the coverability tree by identification of any two vertices representing ω-occurrence sequences that lead to identical markings. Note that, in general, a vertex of the coverability graph can have more than one outgoing edge labeled by the same transition, because different ω-occurrence sequences can lead to the same ω-marking. However, this is not the case for bounded marked nets:

Theorem 27. *The coverability graph and the marking graph of a bounded marked net are identical (up to different co-domains of markings and ω-markings).*

Proof. The result follows immediately from Corollary 26 and the definition of ω-occurrence sequences. □

5 Analysis Techniques

5.1 Behavioral Complexity of Marked Nets

As shown in the previous section, for bounded finite marked nets the marking graph can be constructed to analyze behavioral aspects. For unbounded finite marked nets, the coverability tree provides some information about the behavior. However, even if the marking graph of a marked net is finite, it is not always feasible to explicitly construct this graph because its size can explode with the size of the marked net[9].

Figure 15 shows that the set of reachable markings can grow exponentially with the size of a net: with $n+2$ transitions the marked net has 2^n+1 reachable markings. In the marked net shown in Figure 16, the smallest bound of the place s depends on the initial token count of the place s0: if s0 carries n tokens initially then this bound is 2^n.

Analysis techniques that do not explicitly construct the set of all reachable markings (or ω-markings) have to stick to the structure of the net and its initial marking. Petri net theory provides many such techniques. In this section, we will introduce place and transition invariants, siphons and traps and discuss their relation to deadlock-freedom, liveness, boundedness and reversibility.

Often it is useful to prove that every reachable marking satisfies some given property. The following lemma gives a sufficient condition for this property.[10]

Lemma 28. *A set M of markings of a marked net with initial marking m_0 contains all reachable markings if*

(1) $m_0 \in M$ and
(2) for each marking m of N and each transition t of N, $m \in M$ and $m \xrightarrow{t} m'$ imply $m' \in M$.

[9] This is sometimes called *state explosion problem*, see [Valm98].
[10] This technique is not specific for Petri nets. In general, it is sometimes called *assertional reasoning*.

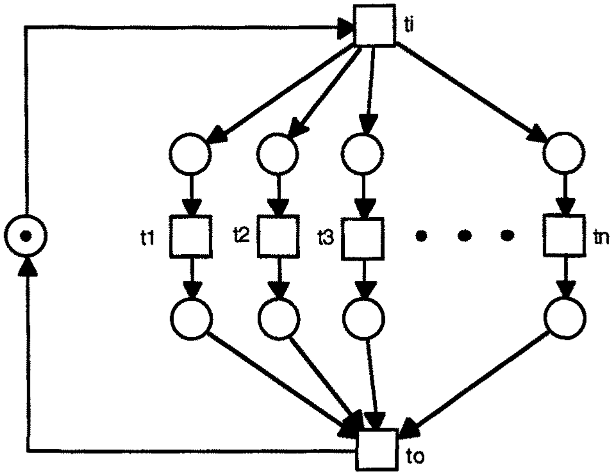


Fig. 15. The set of reachable markings explodes with the size of the net

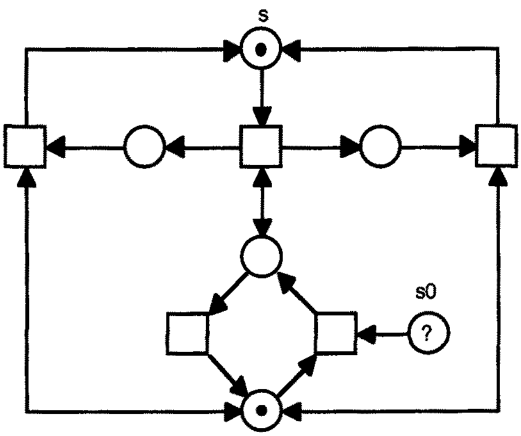


Fig. 16. The set of reachable markings explodes with the number of initial tokens

Proof. Assume that M satisfies conditions (1) and (2). Let m be a reachable marking and let $m_0 \xrightarrow{\sigma} m$, $\sigma = t_1\, t_2\, \ldots t_k$. Assume

$$m_0 \xrightarrow{t_1} m_1 \xrightarrow{t_2} m_2 \xrightarrow{t_3} \cdots \xrightarrow{t_k} m_k.$$

Then $m = m_k$.

We prove that all markings m_i, $0 \leq i \leq k$ are in M, by induction on i.
Base: $i = 0$. Then $m = m_0$. By (1), $m_0 \in M$.

Step: $i > 0$. By the induction hypothesis, $m_{i-1} \in M$. This implies by (2) that $m_i \in M$. □

Notice that conditions (1) and (2) of the above lemma are not necessary for a set M to include all reachable markings. This is because condition (2) applies to all markings m in M and not only to reachable markings in M. Obviously, it would be sufficient to require this condition for reachable markings only. However, the set of reachable markings is unknown in general and hence this modified condition would not be easy to verify.

5.2 Place Invariants

Occurrences of transitions transform the token distribution of a net, but often respect some global properties of markings. For example, the total token count of a set of places remains unchanged if the pre-set and the post-set of the transition contain the same number of places of this set. Place invariants formalize such invariant properties. They moreover allow to weight tokens on places by positive or negative values such that the sum of weighted tokens remains constant.

Given a net N, a *place invariant* is a mapping $i\colon S_N \to \mathbb{Z}$ satisfying $i(s) \neq 0$ for finitely many places and

$$\sum_{s \in {}^\bullet t} i(s) = \sum_{s \in t^\bullet} i(s)$$

for each transition t of N. A place invariant is *nonnegative* if it maps no place to a negative number.

Theorem 29. *If m is a reachable marking of a marked net with initial marking m_0 and i is a place invariant then*

$$\sum_{s \in S_N} i(s) \cdot m(s) = \sum_{s \in S_N} i(s) \cdot m_0(s).$$

Proof. Let M be the set of markings satisfying the above equation. We employ Lemma 28.

Clearly m_0 is in M. Assume a marking m in M and a transition occurrence $m \xrightarrow{t} m'$. Then, for each place s,

$$m'(s) = m(s) - |F \cap \{(s,t)\}| + |F \cap \{(t,s)\}|$$

by the occurrence rule. Hence

$$\sum_{s \in S_N} i(s) \cdot m'(s) = \sum_{s \in S_N} i(s) \cdot m(s) - \sum_{s \in S_N} i(s) \cdot |F \cap \{(s,t)\}| + \sum_{s \in S_N} i(s) \cdot |F \cap \{(t,s)\}|.$$

Observe that $|F \cap \{(s,t)\}| = 1$ if $s \in {}^\bullet t$, and $|F \cap \{(s,t)\}| = 0$ otherwise. Together with the corresponding property for $|F \cap \{(t,s)\}|$ and $t^\bullet$ we obtain

$$\sum_{s \in S_N} i(s) \cdot m'(s) = \sum_{s \in S_N} i(s) \cdot m(s) - \sum_{s \in {}^\bullet t} i(s) + \sum_{s \in t^\bullet} i(s).$$

By definition of place invariants, the last two sums are equal. Hence

$$\sum_{s \in S_N} i(s) \cdot m'(s) = \sum_{s \in S_N} i(s) \cdot m(s).$$

So $m' \in M$.

By Lemma 28, each reachable marking is in M, which implies the result. □

The previous theorem constitutes the salient property of place invariants. For marked nets without dead transitions, the converse direction holds as well:

Theorem 30. *Assume a marked net N without dead transitions. Let m_0 be the initial marking. Let $i: S_N \to \mathbb{Z}$. If each reachable marking m satisfies*

$$\sum_{s \in S_N} i(s) \cdot m(s) = \sum_{s \in S_N} i(s) \cdot m_0(s)$$

then i is a place invariant.

Proof. Let t be a transition. Since t is not dead, there are reachable markings m and m' such that $m \xrightarrow{t} m'$. The assumption implies

$$\sum_{s \in S_N} i(s) \cdot m(s) = \sum_{s \in S_N} i(s) \cdot m'(s).$$

As shown in the proof of Theorem 29,

$$\sum_{s \in S_N} i(s) \cdot m'(s) = \sum_{s \in S_N} i(s) \cdot m(s) - \sum_{s \in {}^\bullet t} i(s) + \sum_{s \in t^\bullet} i(s).$$

Hence

$$\sum_{s \in {}^\bullet t} i(s) = \sum_{s \in t^\bullet} i(s)$$

which finishes the proof. □

Place invariants can be used to prove that a place is bounded:

Theorem 31. *Let s be a place of a marked net N with initial marking m_0. If there is a nonnegative place invariant i satisfying $i(s) \geq 1$ then s is bounded by*

$$\frac{1}{i(s)} \cdot \sum_{s' \in S_N} i(s') \cdot m_0(s').$$

Proof. Since i is nonnegative, $i(s') \geq 0$ for each place s'. Hence

$$i(s) \cdot m(s) \leq \sum_{s' \in S_N} i(s') \cdot m(s').$$

Hence, by Theorem 29, each reachable marking m satisfies

$$i(s) \cdot m(s) \leq \sum_{s' \in S_N} i(s') \cdot m_0(s').$$

Division by $i(s)$ yields the result. □

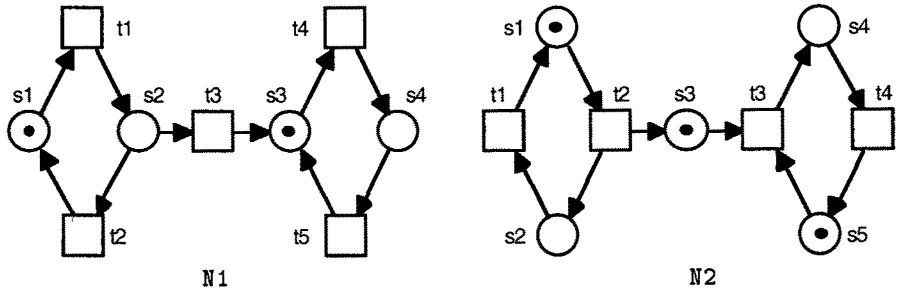


Fig. 17. Two marked nets (copied from Figure 10)

Corollary 32. *A finite marked net is bounded if it has a place invariant i that maps all places to positive numbers.*

Theorem 33. *Let N be a marked net with initial marking m_0 and let i be a nonnegative place invariant. Let S_i be the set of places s satisfying $i(s) > 0$. If $m_0(s) = 0$ for each place in S_i then every transition in ${}^\bullet S_i \cup S_i^\bullet$ is dead at the initial marking.*

Proof. Since i is a place invariant, no reachable marking marks any place in S_i. Hence no reachable marking enables a transition in $S_i^\bullet$. Each transition t in ${}^\bullet S_i$ is also dead because otherwise the occurrence of t would lead to a reachable marking that marks a place in S_i. □

Corollary 34. *If no transition of a marked net is dead at the initial marking then every nonnegative place invariant i either maps all non-isolated places to 0 or satisfies $i(s) \geq 1$ for at least one initially marked non-isolated place.*

This corollary applies in particular to live marked nets.

Figure 17 shows again the two marked nets N1 and N2 from Figure 10. In N1, the set of place invariants is given by all mappings $i: S_{\mathtt{N1}} \to \mathbb{Z}$ satisfying

$$i(\mathtt{s1}) = i(\mathtt{s2}) = i(\mathtt{s3}) = i(\mathtt{s4}).$$

An example is the place invariant that maps each places to 1. Corollary 32 applies and proves all places bounded. Theorem 33 and Corollary 34 do not apply because no transition is dead. Only if the initial marking was the null marking, these results would prove deadness of all transitions.

In N2, the set of place invariants is given by all mappings $i: S_{\mathtt{N2}} \to \mathbb{Z}$ satisfying

$$i(\mathtt{s1}) = i(\mathtt{s2}) \text{ and } i(\mathtt{s3}) = 0 \text{ and } i(\mathtt{s4}) = i(\mathtt{s5}).$$

An example is the place invariant that maps places `s1` and `s2` to 1 and the other places to 0. Using this invariant, Theorem 31 proves that `s1` and `s2` are bounded. If the token on `s1` was missing then Theorem 33 would prove deadness of `t1` and `t2`, using the same place invariant.

5.3 Transition Invariants

Given a net N, a *transition invariant* is a mapping $j: T_N \to \mathbb{Z}$ satisfying $j(t) \neq 0$ for finitely many transitions and

$$\sum_{t \in {}^\bullet s} j(t) = \sum_{t \in s^\bullet} j(t)$$

for each place s of N.

Let σ be a finite sequence of transitions of a net N. The *Parikh mapping* $p_\sigma: T_N \to \mathbb{Z}$ maps each transition t to the number of occurrences of t in σ.

Theorem 35. *If $m \xrightarrow{\sigma} m'$ is a finite occurrence sequence of a net then $m = m'$ if and only if the Parikh mapping p_σ is a transition invariant.*

Proof. Let s be a place. We prove that the equation

$$\sum_{t \in {}^\bullet s} p_\sigma(t) = \sum_{t \in s^\bullet} p_\sigma(t)$$

holds if and only if $m(s) = m'(s)$. The first equation is equivalent to

$$\sum_{t \in {}^\bullet s \setminus s^\bullet} p_\sigma(t) = \sum_{t \in s^\bullet \setminus {}^\bullet s} p_\sigma(t).$$

A token is added to s by a transition occurrence if and only if this transition is in ${}^\bullet s \setminus s^\bullet$. A token is removed if and only if the transition is in $s^\bullet \setminus {}^\bullet s$. Hence the last equation holds if and only if the number of transition occurrences in σ that increase the token count of s coincides with the number of transition occurrences in σ that decrease the token count of s. The latter is exactly the case if $m(s) = m'(s)$. □

Theorem 36. *If the initial marking of a finite bounded marked net enables an infinite occurrence sequence σ then the net has a transition invariant j satisfying*

- $j(t) = 0$ *for each transition t not occurring in σ,*
- $j(t) = 0$ *for each transition t occurring in σ finitely often,*
- $j(t) \geq 1$ *for each transition occurring in σ infinitely often.*

Proof. Let T_ω the set of transitions occurring infinitely often in the sequence σ. Assume $\sigma = t_1\, t_2\, t_3 \dots$ and

$$m_0 \xrightarrow{t_1} m_1 \xrightarrow{t_2} m_2 \xrightarrow{t_3} \cdots.$$

Let k_1 be the first index ($1 \leq k_1$) such that the sequence $t_{k_1+1}\, t_{k_1+2}\, t_{k_1+3} \cdots$ contains only transitions of T_ω.

Define an infinite sequence $k_1\, k_2\, k_3 \ldots$ of indices, $k_1 < k_2 < k_3 < \cdots$, such that, for $i = 1, 2, 3, \ldots$, the transition sequence

$$t_{k_i+1}\, t_{k_i+2}\, t_{k_i+3} \ldots t_{k_{i+1}}$$

contains each transition of T_ω at least once.

Since the marked net is finite and bounded, only a finite set of markings is reachable. Therefore, the markings $m_{k_1}, m_{k_2}, m_{k_3}, \ldots$ cannot be pairwise different. So there are indices i and j, $(1 \leq i < j)$ such that $m_{k_i} = m_{k_j}$. By Theorem 35, the Parikh vector of the transition sequence $t_{k_i+1}\, t_{k_i+2} \ldots t_{k_j}$ is a transition invariant. It is immediate that this transition invariant satisfies the conditions formulated above. □

Corollary 37. *If a finite marked net is live and bounded then it has a transition invariant that maps each transitions to a positive number.*

Proof. By liveness, every transition can always occur again. Therefore there exists an infinite occurrence sequence containing all transitions of the net infinitely often. The result follows by Theorem 36. □

In the example of Figure 17, `N1` is a finite and bounded marked net. It has an infinite occurrence sequence `t4 t5 t4 t5`.... In fact, the net has a transition invariant mapping transitions `t4` and `t5` to 1 and the other transitions to 0.

5.4 Siphons and Traps

Next we consider sets of places, S, that never gain a token once none of their places is marked. As shown in Theorem 33, if a nonnegative place invariant has positive entries only for places of S then no place of S will ever gain a token. Another sufficient condition is that each transition that adds one or more tokens to a place of this set also removes at least one token from some place of the set. This condition is formalized by the notion of a siphon, to be defined next.

A *siphon* is a set S of places satisfying ${}^\bullet S \subseteq S^\bullet$. A siphon is *marked* by a marking m if at least one place of it is marked at m.

Theorem 38. *Assume a marked net with a siphon S. If S is not marked at the initial marking then S is not marked at any reachable marking.*

Proof. We apply Lemma 28 to show that every reachable marking marks no place of S.

Let M be the set of markings that do not mark S. By assumption, the initial marking is in M. Assume a marking m in M and a transition occurrence $m \xrightarrow{t} m'$. Then $t \notin S^\bullet$ because m enables t and m marks no place in S. Since S is a siphon, this implies $t \notin {}^\bullet S$. Hence no place of S can gain a token by the occurrence of t and m' belongs to M, too.

So, by Lemma 28, M includes all reachable markings, which implies the result. □

Consider again the examples shown in Figure 17. The net N1 has the siphons

$$\emptyset, \{\texttt{s1}, \texttt{s2}\}, \{\texttt{s1}, \texttt{s2}, \texttt{s3}, \texttt{s4}\}.$$

After the occurrence of t1 and t3, the siphon $\{\texttt{s1}, \texttt{s2}\}$ is unmarked and can never become marked again.

The siphons of the net N2 are

$$\emptyset, \{\texttt{s1}, \texttt{s2}\}, \{\texttt{s1}, \texttt{s2}, \texttt{s3}\}, \{\texttt{s1}, \texttt{s2}, \texttt{s3}, \texttt{s4}\}, \{\texttt{s1}, \texttt{s2}, \texttt{s3}, \texttt{s4}, \texttt{s5}\}, \{\texttt{s4}, \texttt{s5}\}.$$

Siphons can be used to prove that transitions of a marked net are dead:

Theorem 39. *Assume a marked net with a siphon S which contains at least one non-isolated place. If no transition is dead at the initial marking then some non-isolated place of S is initially marked.*

Proof. Assume that S contains no initially marked non-isolated place. Let s be a non-isolated place in S. By Theorem 38, s is not marked at any reachable marking. Therefore, by Proposition 9 all transition in the set ${}^\bullet s \cup s^\bullet$ are dead. Since s is not isolated, this set is not empty and hence there is an initially dead transition. □

Similarly to siphons, we consider sets of places that never loose all tokens once at least one of their places is marked. A sufficient condition is that each transition that removes at least one token from this set also adds a token. This condition is formalized by the notion of a trap:

A *trap* is a set S of places satisfying $S^\bullet \subseteq {}^\bullet S$. A trap is *marked* by a marking m if at least one place of it is marked at m.

Theorem 40. *Assume a marked net with a trap S. If S is marked at the initial marking then it is marked at every reachable marking.*

Proof. We apply Lemma 28.

Let M be the set of markings of the net that mark at least one place of S. By assumption, the initial marking is in M. Now assume a marking m in M and a transition occurrence $m \xrightarrow{t} m'$. If $t \notin S^\bullet$ then the place of S marked by m remains marked. If $t \in S^\bullet$ then $t \in {}^\bullet S$ because S is a trap. Hence, in this case at least one place in $t^\bullet \cap S$ is marked at m'.

So, by Lemma 28, M includes all reachable markings, which implies the result. □

The traps of the net N1 of Figure 17 are

$$\emptyset, \{\texttt{s3}, \texttt{s4}\}, \{\texttt{s1}, \texttt{s2}, \texttt{s3}, \texttt{s4}\}$$

and the traps of the net N2 are

$$\emptyset, \{\texttt{s1}, \texttt{s2}\}, \{\texttt{s1}, \texttt{s2}, \texttt{s3}, \texttt{s4}, \texttt{s5}\}, \{\texttt{s2}, \texttt{s3}, \texttt{s4}, \texttt{s5}\}, \{\texttt{s3}, \texttt{s4}, \texttt{s5}\}, \{\texttt{s4}, \texttt{s5}\}.$$

By contraposition of the following result, traps can be used to show that a marked net is not reversible:

Theorem 41. *Assume a reversible marked net with a nonempty trap S which contains no isolated place. If no transition is dead at the initial marking then S is initially marked.*

Proof. Let s be a place in S and let t be a transition in ${}^\bullet s \cup s^\bullet$. Since t is not initially dead, there exists a reachable markings m that marks s, by Proposition 9. So the trap S is marked at some reachable marking. Since the initial marking m_0 can be reached from m by reversibility, S is also marked at m_0. □

A first example of this result is given by N2 of Figure 17; this marked net is reversible and has no dead transitions. Hence all nonempty traps are initially marked. As a negative example, consider the marked net N5 of Figure 10. It has no dead transitions. The set $\{\mathtt{s3}, \mathtt{s4}, \mathtt{s5}\}$ is a trap which is not marked initially. Hence the marked net is not reversible.
Finally, siphons and traps provide a sufficient condition for deadlock-freedom:

Theorem 42. *Assume a marked net with at least one transition. If each non-empty siphon without isolated places includes a trap marked at the initial marking then the marked net is deadlock-free.*

Proof. Assume that the marked net is not deadlock-free and let m be a dead reachable marking. Let S be the set of non-isolated places that are not marked at m. We show that S is a non-empty siphon that includes no initially marked trap.

Each transition t is dead at m and hence has an unmarked input place. So $S^\bullet$ contains the set of all transitions. Therefore, ${}^\bullet S \subseteq S^\bullet$. S is not empty because the net has some transition by assumption and S contains a place in the pre-set of this transition. So S is a non-empty siphon without isolated places. By definition, S is not marked at m. Hence, S includes no trap marked at m. Since initially marked traps remain marked, S includes no trap marked at the initial marking. □

Again the nets N1 and N2 of Figure 17 serve as examples. In N1, the siphon $\{\mathtt{s1}, \mathtt{s2}\}$ does not contain any nonempty trap. In N2, the condition formulated in Theorem 42 holds, whence the marked net is deadlock-free.

5.5 Minimal Place Invariants and Minimal Siphons

Theorems 31 and 33 employ nonnegative place invariants. In general, a net possesses infinitely many nonnegative place invariants. We show next that there is a finite set of minimal place invariants such that every nonnegative place invariant can be expressed by a sum of minimal place invariants.

A nonnegative place invariant i is *minimal* if

- $i(s) \geq 1$ for some place s

and there are no two nonnegative place invariants i_1 and i_2 such that

- $i_1(s_1) \geq 1$ for some place s_1,
- $i_2(s_2) \geq 1$ for some place s_2,
- $i(s) = i_1(s) + i_2(s)$ for each place s.

Theorem 43. *The set of minimal place invariants of a finite net is finite.*

Proof. Assume a finite net with an infinite set of minimal place invariants. Let $i_1\, i_2\, i_3 \ldots$ be an infinite sequence of different minimal place invariants. Dickson's Lemma (Lemma 17) implies that this sequence contains place invariants i_j and i_k such that $i_j \neq i_k$ and, for each place s, $i_j(s) \leq i_k(s)$. Consider the mapping $i: S_N \rightarrow \mathbb{Z}$, defined by $i(s) = i_k(s) - i_j(s)$. It is straightforward to verify that i is a nonnegative place invariant, too. Since $i_j \neq i_k$, i_k is not minimal. This contradicts the assumption that the sequence contains only minimal place invariants. □

In Theorems 31 and 33 it suffices to consider minimal place invariants because every place mapped to a positive number by any nonnegative place invariant is mapped to a positive number by a minimal one.

There exists a place invariant i satisfying $i(s) \geq 1$ for each place s if and only if the sum of minimal place invariants enjoys this property. So for Corollary 32 it also suffices to consider minimal place invariants.

The example net **N1** has only one minimal place invariant. This invariant maps all places to 1. The example net **N2** has two minimal place invariants; one of them only maps **s1** and **s2** to 1, the other one only maps **s4** and **s5** to 1 (the other places are mapped to 0).

Clearly, the sets of siphons and traps of a finite net are finite. However, often not every siphon has to be checked for the application of the previous results. In particular for Theorems 39 and 42 it suffices to consider siphons that do not properly include other siphons except the empty set.

A *minimal siphon* S is a nonempty siphon such that no proper subset of S is a nonempty siphon.

If a siphon S contains an isolated place s then $S \setminus \{s\}$ is a siphon, too. So a minimal siphon either contains no isolated places or consists of only one isolated place.

Since the union of traps is again a trap, the *maximal trap* in a given set of places S is the union of all traps included in S. It is the empty set if S includes no nonempty trap. So Theorem 42 can be reformulated as:

Corollary 44. *Assume a marked net with at least one transition and no isolated place. If all maximal traps included in minimal siphons are marked at the initial marking, the marked net is deadlock-free.* □

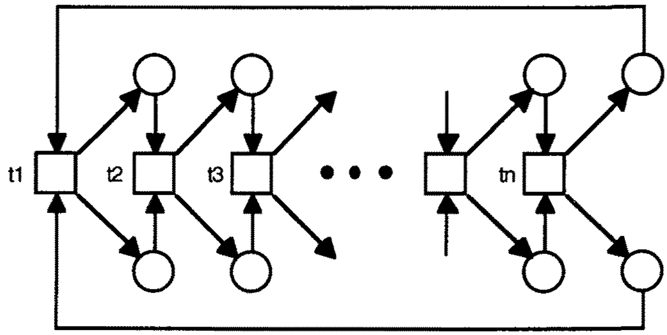


Fig. 18. The sets of minimal siphons, minimal traps and minimal place invariants grows exponentially in the size of the net

The net **N1** has only one minimal siphon, namely {s1, s2}. The maximal trap included in this siphon is empty. The net **N2** has minimal siphons {s1, s2} and {s4, s5}. Both sets are also traps; hence they are the maximal traps included in the siphons.

Minimal place invariants and siphons do not always help to significantly improve the efficiency of analysis techniques. Consider the nets shown in Figure 18. For a fixed number n of transitions, there are 2^n minimal place invariants, mapping one place of each vertical pair of places to 1 and the other place to 0, and there are 2^n minimal siphons, containing exactly one place of each vertical pair each.

6 Concurrent Semantics of Marked Nets

Occurrence sequences provide a *sequential semantics* of marked nets, because they represent each run by a sequence. In general, an occurrence sequence provides little information about dependencies between transition occurrences: if a transition name t' appears after a transition name t in an occurrence sequence then either t' can only occur after t has occurred, or both occurrences are concurrent. The aim of this section is to provide a *concurrent semantics* which explictly represents dependencies between transition occurrences. In particular, this semantics respects concurrency. Relations of this semantics to sequential semantics are studied in depth.

6.1 Causal Nets

Consider again the vending machine from the introduction. Figure 19 shows its representation as a marked net, as in Figure 9. In this example, t5 is only enabled after the occurrence of t3, whereas t3 and t1 can occur concurrently. A *concurrent run* of the transitions t1, t3 and t5 should explicitly represent these dependency relations .

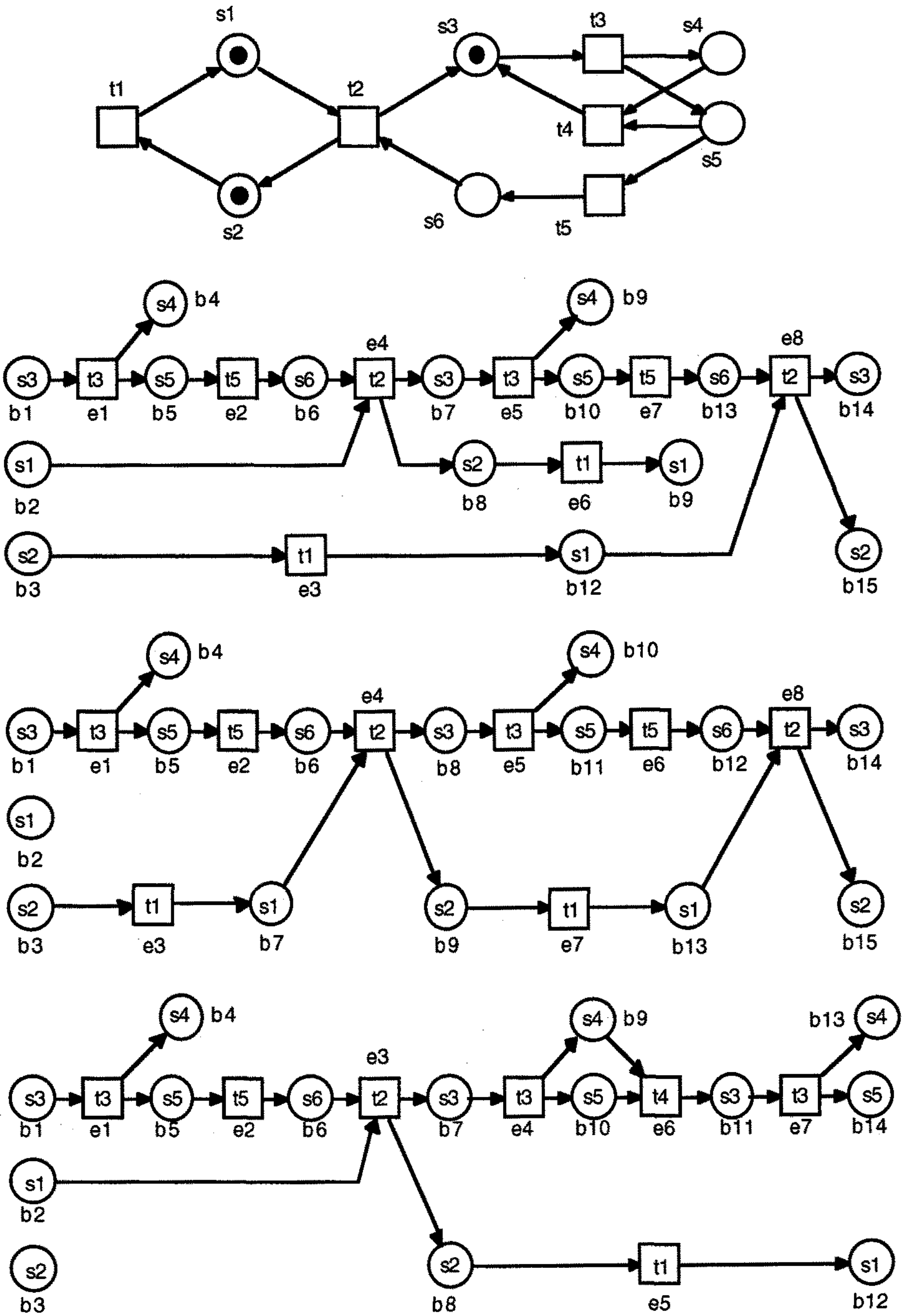


Fig. 19. A marked net with three concurrent runs

The occurrence of `t5` depends on the previous occurrence of `t3` because `t3` produces a token that is consumed by `t5`. Transition occurrences together with produced and consumed tokens and their mutual relationship are most conveniently represented by a special kind of labeled nets. Thus each concurrent run of a marked net is represented by a net, called causal net. To avoid confusion, the marked net will called *system net* in the sequel.

Figure 19 shows a marked system net and three causal nets representing concurrent runs. Each place of a causal net represents a single token occurrence in a system net. We call places of causal nets *conditions*, emphasizing their role as pre- and post-conditions of transition occurrences. Similarly, each transition of a causal net represents a single occurrence of a transition of a system net. Transitions of causal nets are called *events*.

The elements of the causal nets are usually written `b1`, `b2`, ... (conditions)[11] and `e1`, `e2`, ... (events). Moreover, each element is inscribed by the corresponding element of the system net: an event inscribed by t represents an occurrence of t, and a condition in the post-set of the event stands for a token produced by the transition occurrence. The condition inscription denotes the place of the system net, to which the token is added. The initial token distribution is reflected by according conditions of a causal net with empty pre-set. The causal nets of Figure 19 are finite, but causal nets may be infinite, thus representing infinite runs.

According to the vending machine interpretation of the example net, the first two causal nets represent concurrent runs where two coins are inserted and accepted, and two items are dispensed. In the first concurrent run, the first dispensed item is the one initially in the storage. The second item had to be refilled before it is dispensed. In the second run, the item initially in the storage is not touched. Instead both diepensed items depend on preceding occurrences of the refill transition `t1`. Concurrent semantics distinguishes these two runs because the mutual dependencies of transition occurrences are different. One could argue that the second run is more efficient because the occurrences of `t2` do not unnecessarily wait. Moreover, if items should not stay too long in the storage (such as warm beverages), the second run is not desirable. The third run is different. Here the second inserted coin is rejected. In contrast to the other runs, some condition inscribed by `s5` is followed by an event inscribed by `t4`. Thus, in this run a token an `s5` is removed by the occurrence of `t4`.

Formally, a *causal net* is a net with conditions (places) and events (transitions) satisfying the following properties:

(1) the flow relation is acyclic, i.e.
no path with at least two elements leads from an element to itself;
(2) conditions are not branched, i.e. each condition b satisfies $|{}^\bullet b| \leq 1$, $|b^\bullet| \leq 1$;
(3) only finitely many conditions b have an empty pre-set ${}^\bullet b$;
(4) for each event e, both sets ${}^\bullet e$ and $e^\bullet$ are finite and nonempty;
(5) for each element x, only finitely many different paths lead to x.

[11] from the German *Bedingungen*

For a given causal net K[12], the symbols B_K, E_K and F_K denote its sets of conditions, events and directed arcs, respectively. Since the flow relation of a causal net K is acyclic by definition, it induces a partial order $\preceq_K$ between the elements of K: $x \preceq_K y$ if and only if there is a path leading from x to y. We say that x *precedes* y in K if $x \neq y$ and $x \preceq_K y$. According to (4), each element is preceded by finitely many elements.

A *line* of a causal net K is a maximal set of elements that are pairwise ordered by $\preceq_K$ (maximality w.r.t. set inclusion).

In the first causal net of Figure 19, the set {b2, e4, b8, e6, b9} is a line.

Proposition 45. *For each path of a causal net, all elements appearing in the path belong to a single line.* □

A *cut* is a maximal set of elements such that every two distinct elements of the set are not ordered by $\preceq_K$.

In the first causal net of Figure 19, the set {b9, e7, b11, b12} is a cut.

The maximality of a cut implies the following proposition:

Proposition 46. *Given a cut X of a causal net K, each element y of K either belongs to X or precedes an element of X or is preceded by an element of X.* □

6.2 Causal Nets with Canonical Marking

Representing a single concurrent run of a system net, there is no need for markings of causal nets. Markings for causal nets are nevertheless useful for technical reasons. A *canonical* initial marking of a causal net K assigns to each minimal (w.r.t. $\preceq_K$) condition one token and no token to all other conditions. As will turn out, occurrence sequences of system nets and occurrence sequences of corresponding causal nets are closely related. This relation demonstrates that the concurrent semantics provided by causal net respects the sequential semantics provided by occurrence sequences.

For each causal net K, we call the marking l_0 defined by

$$l_0(b) = \begin{cases} 1 & \text{if } {}^\bullet b = \emptyset \\ 0 & \text{if } {}^\bullet b \neq \emptyset \end{cases}$$

the *canonical initial marking* of K.

In all causal nets of Figure 19, the canonical initial markings assign one token to the conditions b1, b2 and b3 and no token to any other condition.

[12] from the German *Kausalnetz*

Theorem 47. *Assume a causal net K with canonical initial marking l_0. A marking l of K is reachable from l_0 if and only if*

- *$l(b) \leq 1$ for each condition b, and*
- *the set $\{b \in B_K \mid l(b) = 1\}$ is a finite cut of K.*

Proof.
($\Longleftarrow$) Let B denote the finite cut $\{b \in B_K \mid l(b) = 1\}$. Let E be the set of events preceding conditions of B. By the definition of causal nets, E is finite because B is finite. We proceed by induction on $|E|$.
Base: $k = 0$. Then all conditions in B have an empty pre-set. Since B is a cut, B contains all conditions with empty pre-set. Therefore, $l = l_0$.
Step: $k > 0$. Assume a maximal event e in E (maximality w.r.t. $\preceq_K$).

We claim that $e^\bullet \subseteq B$. Let b in $e^\bullet$.

By the maximality of e, either $b^\bullet = \emptyset$ or $b^\bullet = \{e'\}$ for some event e' that does not belong to E. In both cases, b does not precede any condition of B. Since b is not branched, every condition preceding b also precedes e. Since e precedes some condition of B and conditions of B are not ordered, no condition preceding b belongs to B. Proposition 46 implies $b \in B$.

Define $B' = (B \setminus e^\bullet) \cup {}^\bullet e$. B' is finite because B as well as ${}^\bullet e$ are finite. We claim that B' is a cut. To this end, we first show that no condition in ${}^\bullet e$ is ordered with any other condition in B'. Then we prove that every element of K either belongs to B' or is ordered with at least one condition of B'.

Conditions in ${}^\bullet e$ are mutually not ordered because conditions are not branched and e does not precede a condition in ${}^\bullet e$ because K is acyclic. There is no path from a condition in ${}^\bullet e$ to a condition in $B \setminus e^\bullet$ because this path would pass through a condition in $e^\bullet$ but $e^\bullet \subseteq B$ and conditions in B are not ordered. There is no path from a condition in $B \setminus e^\bullet$ to a condition in ${}^\bullet e$ because this path could be extended by e and a condition in $e^\bullet$; since $e^\bullet \subseteq B$, this extended path would establish an order between distinct conditions of B.

Let x be an element of K. By Proposition 46, either x precedes some b in B, or some b in B precedes x, or x is equal to some b in B. If this condition b can be taken from $B \setminus e^\bullet$ then x is ordered with an element of B', too. Otherwise $b \in e^\bullet$. If $x \in e^\bullet$ or $x = e$ then any element of ${}^\bullet e$ precedes x. If $x \in {}^\bullet e$ then $x \in B'$. In all other cases, x precedes b if and only if x precedes a condition of ${}^\bullet e$ and b precedes x if and only if a condition of ${}^\bullet e$ precedes x. Since ${}^\bullet e \subseteq B'$, x is ordered with some condition of B'. This finishes the proof of the claim that B' is a cut.

Since the set of events preceding B' is $E \setminus \{e\}$, the induction hypothesis applies to B'. Hence some reachable marking l' marks exactly the conditions of B', by one token each. The event e is enabled at l'. Its occurrence leads to l.

($\Longrightarrow$) Let B be the set of conditions of some line X of K. We claim that, for any reachable marking, the number of all tokens on conditions of B is one. The claim is proven by help of Lemma 28. Let L be the set of markings satisfying the above property.

The line X contains at least one element with empty pre-set, because every element has a finite past. It contains at most one such element because every two elements of X are connected by a path. This unique minimal element of X is a condition, because events have non-empty pre-sets. By the definition of l_0, this condition is initially marked with one token, and all other conditions of X are initially unmarked. Hence l_0 belongs to L.

Assume a marking l in L and an event occurrence $l \xrightarrow{e} l'$. The token distribution on B is only changed if e belongs to the line X, because conditions are not branched. In this case, B contains exactly one condition in ${}^\bullet e$ and exactly one condition in $e^\bullet$ because e has nonempty pre- and post-sets. So the occurrence of e does not change the number of tokens on places in B. Therefore, l' also belongs to L, which finishes the proof of the claim.

Let l be a reachable marking. Since each condition is an element of a line, the claim implies that no condition carries more than one token at l, and that each condition is ordered with some condition marked at l. Proposition 45 implies that conditions marked at l are mutually not ordered. So the set of conditions marked at l constitutes a cut. This cut is finite because initially a finite set of conditions is marked and each event has a finite post-set. □

Corollary 48. *Each causal net marked with its canonical initial marking is a 1-bounded marked net.* □

The next result implies that occurence sequences of a causal net K respect the order $\preceq_K$.

Theorem 49. *Assume a causal net K with events e and e' satisfying $e \preceq_K e'$ (including the case $e = e'$). The canonical initial marking enables no occurrence sequence $\sigma_1\, e'\, \sigma_2\, e$.*

Proof. After the occurrence of e', the event e' precedes some marked condition. This situation is not changed by the occurrence of other events because all events have nonempty post-sets. Let l be a marking reached after an occurrence sequence $\sigma_1\, e'\sigma_2$. Since $e \preceq_K e'$, the event e precedes a condition marked at l. By Theorem 47, no condition in ${}^\bullet e$ is marked at l. Since ${}^\bullet e \neq \emptyset$, the event e is not enabled at l. □

Corollary 50. *No event of a causal net occurs more than once in an occurrence sequence enabled at the canonical initial marking.* □

Finally, the following theorem implies that each event of an occurrence sequence can eventually occur.

Theorem 51. *Let $e_1 e_2 e_3 \ldots$ be a sequence of events of a causal net K containing each event of E_K exactly once such that $e_i \preceq_K e_j$ implies $i \leq j$. Then this sequence is enabled at the canonical initial marking.*

Proof. By Proposition 3, it suffices to consider finite prefixes $e_1\, e_2 \ldots e_k$. The proof is by induction on k:
Base: $k = 0$. The empty sequence is enabled at any marking.
Step: $k > 0$. By the induction hypothesis, the canonical initial marking enables the occurrence sequence $e_1\, e_2 \ldots e_{k-1}$. Assume this sequence leads to l_{k-1}. Each event preceding e_k has occurred in this sequence. So each condition in ${}^\bullet e_k$ is marked at l_{k-1}. Hence l_{k-1} enables e_k. □

6.3 Process Nets

Since a causal net K represents a concurrent run of a marked net N, its conditions correspond to places of N and its events correspond to transitions of N. This relation is formalized by a *labeling function* $\pi\colon (B_K \cup E_K) \to (S_N \cup T_N)$.

A labeled causal net K is a *process net* of a net N with initial marking m_0 if the labeling function π enjoys the following properties:

(6) for each condition b in B_K, $\pi(b) \in S_N$;
(7) for each event e in E_K, $\pi(e) \in T_N$;
(8) for each event e in E_K, π generates bijections

$$\pi_e^-\colon {}^\bullet e \to {}^\bullet(\pi(e)) \;\text{ and }\; \pi_e^+\colon e^\bullet \to (\pi(e))^\bullet,$$

defined by $\pi_e^-(b) = \pi(b)$ for each b in ${}^\bullet e$ and $\pi_e^+(b) = \pi(b)$ for each b in $e^\bullet$;
(9) for each place s in S_N, $m_0(s) = |\{b \in B_K \mid {}^\bullet b = \emptyset \wedge \pi(b) = s\}|$.

Graphically, we denote the labels as inscriptions in the elements of a causal net.

The following two theorems tightly relate the occurrence sequences of a marked net and the occurrence sequences of its process nets.

Theorem 52. *Assume a marked net N with initial marking m_0 and a process net K with labeling function π. Let l_0 be the canonical initial marking of K and assume a finite occurrence sequence $e_1\, e_2 \ldots e_k$ of K, leading from l_0 to a marking l_k. Then $\pi(e_1)\, \pi(e_2) \ldots \pi(e_k)$ is an occurrence sequence of N, enabled at m_0. It leads to the marking m_k, satisfying for each place s the equation*

$$m_k(s) = |\{b \in B_K \mid l_k(b) = 1 \wedge \pi(b) = s\}|.$$

Proof. The proof is by induction on k.
Base: $k = 0$. The empty sequence is enabled by any marking. The equation for m_k coincides with the equation for m_0 in the definition of process nets.
Step: $k > 0$. Let $l_0 \xrightarrow{e_1 \ldots e_{k-1}} l_{k-1}$. By the induction hypothesis, m_0 enables the occurrence sequence $\pi(e_1)\pi(e_2) \ldots \pi(e_{k-1})$. Assume this sequence leads to m_{k-1}. Since e_k is enabled at l_{k-1}, all conditions in ${}^\bullet e_k$ are marked at l_{k-1}. By definition of a process net, π generates a bijection between these conditions and the places in ${}^\bullet(\pi(e_k))$.

Also by the induction hypothesis,

$$m_{k-1}(s) = |\{b \in B_K \mid l_{k-1}(b) = 1 \wedge \pi(b) = s\}|$$

for each place s in S_N. Hence each place in $^\bullet(\pi(e_k))$ is marked at m_{k-1}. So $\pi(e_k)$ is enabled at m_{k-1}. Since, for each place s in $^\bullet(\pi(e_k))$, e_k removes one token from a condition labeled by s, and, for each place s in $(\pi(e_k))^\bullet$, e_k adds one token to a condition labeled by s, we obtain

$$m_k(s) = |\{b \in B_K \mid l_k(b) = 1 \wedge \pi(b) = s\}|$$

for each place s in S_N. □

Proposition 2 implies:

Corollary 53. *Assume a marked net N with initial marking m_0 and a process net K with labeling function π. Let l_0 be the canonical initial marking of K and assume an infinite occurrence sequence $e_1\, e_2\, e_3 \ldots$ of K, enabled by l_0. Then $\pi(e_1)\, \pi(e_2)\, \pi(e_3) \ldots$ is an occurrence sequence of N, enabled at m_0.* □

For the converse direction we stick to marked nets satisfying the following property:

> **(N1)** *Both the pre-set and the post-set of each transition is nonempty and finite. The set of initially marked places is finite.*

By definition of a process net, occurrences of transitions with empty or infinite pre- or post-set cannot be represented in a causal net. If infinitely many places are initially marked then there is no corresponding finite initial cut of a process net.

Theorem 54. *Assume a marked net N with initial marking m_0 satisfying property* **(N1)**. *For each occurrence sequence $t_1\, t_2 \ldots t_k$, there exists a process net K with labeling function π such that its canonical initial marking l_0 enables an occurrence sequence $e_1\, e_2 \ldots e_k$ satisfying $\pi(e_i) = t_i$ for $1 \le i \le k$.*

Proof. The proof is by induction on k.
Base: $k = 0$. We only have to prove the existence of a process net K with appropriate labeling function π. To this end, a causal net without events and arcs is constructed: for each place s of N, the causal net contains $m_0(s)$ conditions labeled by s. It is easily verified that this net satisfies the defining properties of causal nets and process nets.
Step: $k > 0$. By the induction hypothesis there exists a process net K' with labeling function π', canonical initial marking l'_0, and an occurrence sequence

$$l'_0 \xrightarrow{e_1 \ldots e_{k-1}} l'_{k-1} \text{ such that } \pi(e_i) = t_i \text{ for } 1 \le i \le k-1.$$

Without loss of generality assume that every event of K occurs in this occurrence sequence. By Corollary 50, no event occurs more than once. So l'_{k-1} only marks conditions with empty post-set.

We construct the causal net K by adding to K' a new event e_k and, for each place s_i in $t_k^\bullet$, a new condition b_i. The labeling function π extends π' by $\pi(e_k) = t_k$ and $\pi(b_i) = s_i$ for each s_i in $t_k^\bullet$.

By Theorem 52,

$$m_{k-1}(s) = |\{b \in B_K \mid l'_{k-1}(b) = 1 \wedge \pi(b) = s\}|$$

for each place s. The marking m_{k-1} enables the transition t_k. Hence, for each place s_i in $^\bullet t_k$, at least one condition b_i labeled by s_i is marked at l'_{k-1} (there might be several suitable conditions). For each place s_i in $^\bullet t_k$, choose one of the conditions b_i labeled by s_i and add a directed arc from this condition to e_k. Finally, add directed arcs from e_k to all new conditions.

This construction obviously yields a causal net. The labeling function π satisfies the conditions formulated for process nets, because π' satisfies these conditions and the vicinity of e_k is respected by π. Since the new conditions are not marked initially, the occurrence sequence $e_1 \ldots e_{k-1}$ leads from the canonical initial marking of K to a marking l_{k-1} that coincides with l'_{k-1} on all conditions and maps the new conditions to 0. By construction, the event e_k is enabled at l_{k-1}. Thus, $e_1\, e_2 \ldots e_k$ is an occurrence sequence enabled at l_0. □

The construction of the causal net in the previous proof is not unique. However, if the marked net is 1-bounded then no reachable marking of the causal net marks two equally labeled conditions. Hence, for 1-bounded marked nets there is always exactly one way to connect marked conditions to a new event. Therefore, for 1-bounded marked nets, each occurrence sequence corresponds uniquely to a process net with a corresponding occurrence sequence. Figure 20 shows that the same does not hold for 2-bounded marked nets.

Consider occurrence sequences σ and σt of a marked net. Then the construction in the previous proof yields causal nets K and K' such that $B_K \subseteq B_{K'}$, $E_K \subseteq E_{K'}$ and $F_K \subseteq F_{K'}$. Moreover, the occurrence sequence of K is a prefix of the occurrence sequence of K'. Hence, for each infinite occurrence sequence σ of a marked net, the union of the respective sets of conditions, events, and arcs yields an infinite causal net. This infinite causal net has an infinite occurrence sequence such that the sequence of labels of its events is σ. So Theorem 54 implies:

Corollary 55. *Assume a marked net N with initial marking m_0 satisfying the conditions of* **(N1)**. *For each infinite occurrence sequence $t_1\, t_2\, t_3 \ldots$, there exists a process net K with labeling function π such that its canonical initial marking l_0 enables an occurrence sequence $e_1\, e_2\, e_3 \ldots$, satisfying $\pi(e_i) = t_i$ for $1 \leq i$.* □

Theorem 47 and Theorem 51, together with the previous results imply the following corollaries:

Corollary 56. *Assume a marked net N with initial marking m_0 satisfying the conditions of* (**N1**). *A marking m is reachable from m_0 if and only if there is a process net K with labeling function π such that some finite cut B of K satisfies $B \subseteq B_K$ and, for each place s of N, $m(s) = |\{b \in B | \pi(b) = s\}|$.* □

The set of events E_K of a causal net K is partially ordered by $\preceq_K$. A (finite or infinite) sequence $e_1\, e_2\, e_3 \ldots$ of events is a *sequentialization* of E_K if

- every event of E_K appears exactly once in the sequence, and
- for each two events e and e' satisfying $e \preceq_K e'$, e appears before e' in the sequence.

Corollary 57. *Assume a marked net N with initial marking m_0 satisfying the conditions of* (**N1**). *A sequence $e_1\, e_2\, e_3 \ldots$ is a sequentialization of the set of events of a process net with labeling function π if and only if $\pi(e_1)\,\pi(e_2)\,\pi(e_3) \ldots$ is an occurrence sequence of N, enabled at m_0.* □

6.4 Process Nets and Boundedness

Behavioral properties of a marked net based on reachable markings and transition occurrences can be reformulated in terms of process nets. For example, reversibility requires that for each finite process net leading to a cut X, there is another process net leading from X to a cut that corresponds to the initial marking. Boundedness can be characterized in a more elegant way:

Theorem 58. *A finite marked net N with initial marking m_0 satisfying the conditions of* (**N1**) *is unbounded if and only if some process net has an infinite cut.*

Proof.
($\Longleftarrow$) Assume a process net K with an infinite cut B. If B contains an event e then $(B \setminus \{e\}) \cup e^\bullet$ is also an infinite cut because events have nonempty post-sets and conditions are not branched. So we can assume without loss of generality that B contains only conditions.

We proceed indirectly and assume that N is bounded. Since N is finite, there is an upper bound k for the total number of tokens on all places. Consider an arbitrary finite subset B' of B with more than k conditions.

Let E be the set of all events preceding a condition of B'. Since B is a cut, no condition of B' precedes an event of E. By definition of a causal net, E is finite, because B' is finite. Since for each e in E, all events preceding E are contained in E as well, the canonical initial marking of K enables an occurrence sequence containing exactly the events of E. The marking reached by this sequence marks at least all conditions of B'. By Corollary 56, there is a corresponding reachable marking of the marked net such that the total token count of all places at this marking exceeds k, which contradicts the choice of k.

($\Longrightarrow$) Since the marked net is not bounded, some place is not bounded. By Theorem 23, there is a minimal ω-occurrence sequence σ leading to an ω-marking $\overline{m}$ such that $\overline{m}(s) = \omega$ for some place s. By definition of ω-occurrence sequences, $\sigma = \sigma_1\, \sigma_2$ such that, with

$$m_0 \xrightarrow{\sigma_1} m_1 \xrightarrow{\sigma_2} m_2,$$

each place s' satisfies $m_2(s') \geq m_1(s')$ and the place s satisfies $m_2(s) > m_1(s)$.

Now modify the net N as follows: Add a new transition $\overline{t}$, a new place $\overline{s}$, and new arcs $(s, \overline{t})$ and $(\overline{t}, \overline{s})$. This modified net will be called $\overline{N}$. Let $\overline{m}_0$ be the initial marking of $\overline{N}$ which coincides with m_0 on all places of N and assigns no token to $\overline{s}$.

Since the number of tokens on s increases by the occurrence of σ_2, there is an occurrence sequence

$$\overline{m}_0 \xrightarrow{\sigma_1} \overline{m}_1 \xrightarrow{\sigma_2\, \overline{t}} \overline{m}_2.$$

Then $\overline{m}_2(s') \geq \overline{m}_1(s')$ for each place s' of $\overline{N}$. So the sequence $\sigma_2 \overline{t}$ can be iterated any number of times:

$$\overline{m}_0 \xrightarrow{\sigma_1} \overline{m}_1 \xrightarrow{\sigma_2\, \overline{t}} \overline{m}_2 \xrightarrow{\sigma_2\, \overline{t}} \overline{m}_3 \xrightarrow{\sigma_2\, \overline{t}} \cdots .$$

By Theorem 54, there exists a process net K with labeling function π such that its canonical initial marking enables an infinite occurrence sequence of events where the sequence of event labels is

$$\sigma_1\, \sigma_2\, \overline{t}\, \sigma_2\, \overline{t}\, \sigma_2\, \overline{t} \ldots .$$

This causal net contains infinitely many events labeled by $\overline{t}$. Let B be the set of conditions in the pre-set of these events. Then B is also infinite.

Now consider the causal net gained from the previous one by deleting all events labeled by $\overline{t}$, all conditions labeled by $\overline{s}$, and all adjacent arcs. Then all conditions in B have an empty post-set. Hence they are mutually not ordered. It is easy to see that, taking the same labels as before, this causal net constitutes a process net of the original marked net. Clearly, the set B is again an infinite set of mutually not ordered conditions.

It remains to prove that B is included in an infinite cut. Let E be the set of all events preceding a condition of B. Let B' be the set of all minimal conditions (w.r.t. $\preceq_K$) not preceding any event in E. By minimality, conditions in B' are mutually not ordered. B' includes B because each condition in B either has an empty pre-set or the pre-set consists of an event in E. Finally, B' is maximal (w.r.t. set inclusion) because each minimal element not preceding a condition of B does not precede an event of E and hence belongs to B'. So B' is an infinite cut. □

Finally, the following proof of the Strongly-Connectedness-Theorem (Theorem 11) should help to demonstrate the advantages of causal semantics. Here we stick to nets satisfying the conditions of **(N1)**.

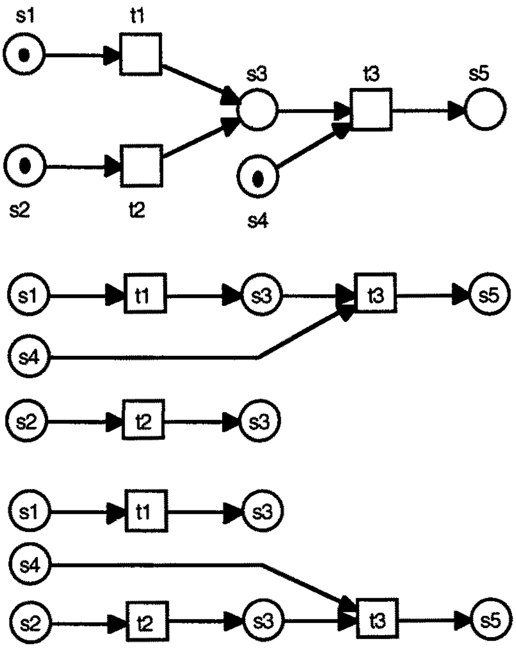


Fig. 20. A marked net and two of itsprocesses

Theorem 59. *Every weakly connected live and bounded marked net N satisfying the conditions of* (**N1**) *is strongly connected.*

Proof. By Proposition 1 it suffices to prove that, for each arc (x, y), there is a directed path leading from y to x.

Let t be the transition in $\{x, y\}$ and let s be the place in $\{x, y\}$. By liveness, there is an occurrence sequence with infinitely many occurrences of t. Hence, some process has infinitely many events labeled by t. Each of these events has a condition labeled by s in its pre-set or post-set. Let B be the infinite set of these conditions. By boundedness, the process net has no infinite cut. Hence there is a path connecting two distinct conditions of B. Since conditions are not branched, this path contains an event labeled by t.

By the definition of process nets, the sequence of labels of the elements of a path of a process net is a path of N. So some path of N leads from s via t back to s. In particular, some path leads from y to x. □

6.5 Expressive Power of Process Nets

The previous results have proven a close relationship between the sequential semantics of a marked net given by its occurrence sequences and the concurrent semantics given by process nets. Roughly speaking, an occurrence sequence determines a totally ordered set of transition occurrences. This order respects causal dependencies of transitions occurrences, but also concurrent transition occurrences are ordered.

In contrast, process nets reflect exactly the causal dependencies. The example of Figure 20 shows that causal dependencies of transition occurrences cannot be obtained from corresponding occurrence sequences. The figure shows a marked net and two of its process nets (labels are again depicted by inscriptions in conditions and events). In each process net, every transition of the marked net occurs exactly once. The first process net represents a concurrent run where the transition `t3` causally depends on the previous occurrence of `t1`. In the second process net, transition `t3` causally depends on the previous occurrence of `t2`. The first process net corresponds to the occurrence sequences

`t1 t2 t3` , `t1 t3 t2` and `t2 t1 t3`.

The second process net corresponds to the occurrence sequences

`t1 t2 t3` , `t2 t1 t3` and `t2 t3 t1`.

The occurrence sequences `t1t3t2` and `t2t3t1` belong to only one causal net each. The other two occurrence sequences, however, do not carry sufficient information to decide which concurrent run they respectively represent.

The reader is recommended to prove that the first two process nets of Figure 19 also have a common occurrence sequence.

7 Arc-Weights, Capacities, and Inhibitor Arcs

7.1 Arc-Weights

A *weighted arc* specifies that more than one token is removed from a place or added to a place by a single occurrence of a transition. As an example, consider the vending machine of the introduction, shown in Figure 5. If the arcs

(`item storage`, `dispense item`) and (`dispense item`, `request for refill`)

are both weighted by 2 then each occurrence of `dispense item` removes two tokens from `item storage` and adds two tokens to `request for refill`. Thus, this modified vending machine sells items in pairs. Similarly, if the arcs

(`refill`, `item storage`) and (`item storage`, `dispense item`)

are both weighted by 4, four items are sold by each occurrence of `dispense item`. In this case, a token on `request for refill` indicates that the storage is empty, whereas no token means that the storage carries four items.

The formal definition of a net N with arc-weights replaces the flow relation F_N by a weight function shaped

$$w_N: (S_N \times T_N) \cup (T_N \times S_N) \to \mathbb{N}.$$

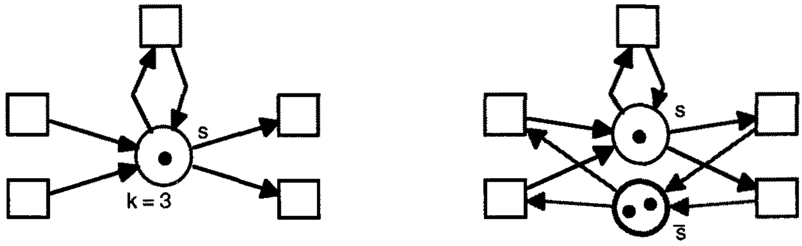


Fig. 22. A place with a weak capacity and a corresponding complement place

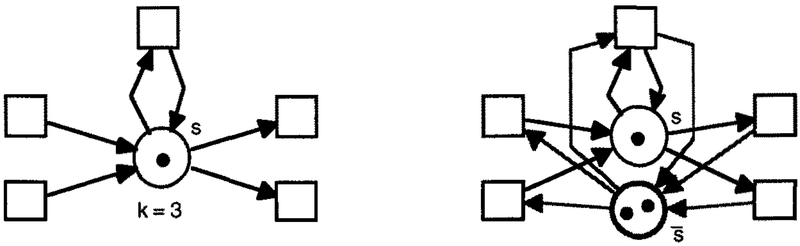


Fig. 23. A place with a strong capacity and a corresponding complement place

7.2 Capacities

A *capacity restriction* (*capacity*, for short) specifies that the token count of a place should never exceed a previously fixed value. Consider again the example of Figure 5. If the place **item storage** has a specified capacity 4, then no reachable marking should assign more than four tokens to the place. Since the place is 4-bounded, this capacity is respected.

Now assume the same marked net, but without the place **request for refill** and adjacent arcs. In this net, the transition **refill** always can occur, whence the place **item storage** is unbounded. If this place is given a capacity 4 then the behavior of the marked net with capacity restriction equals the behavior of the original marked net, i.e. both marked nets have the same occurrence sequences.

Formally, capacities are represented by assigning each place an element of $\{\omega, 1, 2, 3, \ldots\}$. The value ω means that there is no capacity restriction for this place. The initial marking has to respect all capacities. A modified occurrence rule prevents each transition occurrence that would violate any capacity restriction. Instead of presenting this modified model we will show how capacity restrictions are equivalently replaced by *complement places*. We distinguish *weak capacities* and *strong capacities*.

Figure 22 shows how to implement a weak capacity restriction of a place by means of a complement place. Consider a place s with a weak capacity k, $k \neq \omega$. The complement of s is a new place $\overline{s}$ (the shaded place in the figure) satisfying

$$^\bullet\overline{s} = s^\bullet \setminus {}^\bullet s \text{ and } \overline{s}^\bullet = {}^\bullet s \setminus s^\bullet.$$

This complement place $\overline{s}$ is initially marked by $k - m_0(s)$, where $m_0(s)$ is the

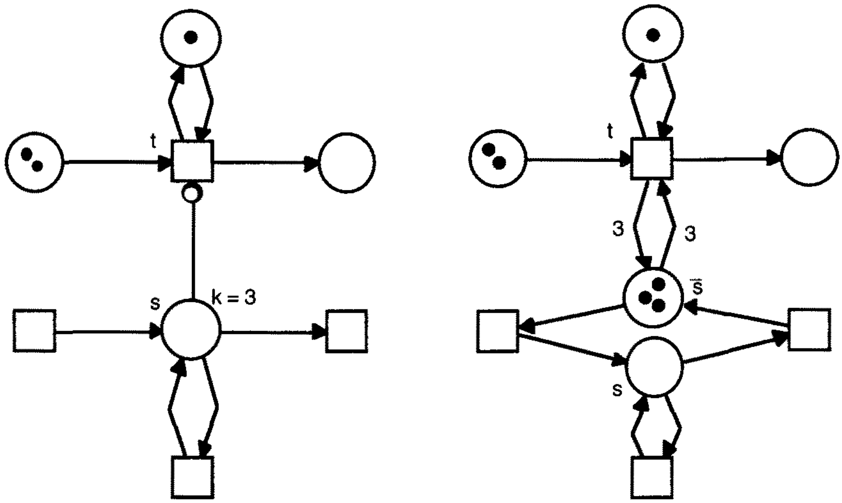


Fig. 24. A net with an inhibitor arc

initial marking of s. By construction, the total token count of s and $\overline{s}$ is equal to k for each reachable marking. Since markings are nonnegative, this construction guarantees that s is bounded by k.

A *strong* capacity differs from a weak capacity only in the presence of *self-loops*, i.e. of mutually connected elements. Given a strong capacity k, $k \neq \omega$ of a place s, not only the token count on s never exceeds k but moreover a transition t in ${}^\bullet s$ is only enabled at a marking m if $m(s) < k$. Figure 23 shows how to implement a strong capacity restriction of a place by means of a complement place. The complement of a place s with strong capacity k, $k \neq \omega$ is a new place $\overline{s}$ satisfying

$$ {}^\bullet\overline{s} = s^\bullet \quad \text{and} \quad \overline{s}^\bullet = {}^\bullet s. $$

Again, this place is initially marked by the value $k - m_0(s)$, and by construction s is k-bounded. Moreover, for each reachable marking m, $m(s) = k$ implies $m(\overline{s}) = 0$. Since a transition t in ${}^\bullet s$ is in $\overline{s}^\bullet$, t can only occur when s carries less than k tokens.

Strong capacities establish a close relationship between marked nets and elementary net systems: Each elementary net system can be considered a marked net with strong capacity 1 for all places, and vice versa.

7.3 Inhibitor Arcs

An *inhibitor arc* connecting a place and a transition specifies that the transition should only occur at a marking if the place is unmarked. Formally, inhibitor arcs of a net N constitute a subset of $S_N \times T_N$. The enabling condition of the occurrence rule is extended by the above constraint.

Graphically, an inhibitor arc is represented by a special arrow head. Figure 24 shows on the left hand side a marked net with inhibitor arc $(\mathbf{s}, \mathbf{t})$.

It is well known that inhibitor arcs can properly increase the expressive power of a net. The argument is that *two-counter machines* can be modeled by deterministic finite nets with (at least two) unbounded places and adjacent inhibitor arcs. Since two-counter machines can simulate Turing machines, it is not decidable if they ever halt. The two-counter machine halts if and only if the corresponding marked net is not deadlock-free, because this model has only one maximal occurrence sequence. However, deadlock-freedom for marked nets without inhibitor arcs easily reduces to the reachability problem [EsNi94], which is decidable [Mayr84].

Now consider inhibitor arcs connected to bounded places. Each b-bounded place s can be assigned the weak capacity b without changing its behavior. As shown above, a capacity restriction can equivalently be replaced by a complement place $\overline{s}$. Since, for each reachable marking, s is unmarked if and only if $\overline{s}$ carries b tokens, the inhibitor arc is equivalently replaced by conventional arcs from and to $\overline{s}$, weighted by b. Figure 24 shows an example.

References

[BeDe90] E. Best and J. Desel. Partial order behaviour and structure of Petri nets. Formal Aspects of Computing 2(2):123–138, 1990.

[BeFe88] E. Best and C. Fernandez C. Nonsequential Processes – a Petri Net View. EATCS Monographs on Theoretical Computer Science Vol. 13, Springer-Verlag, 1988.

[Brau80] W. Brauer (Ed.) Net Theory and Applications. Vol. 84 of Lecture Notes in Computer Science, Springer-Verlag, 1980.

[BRR87] 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, Springer-Verlag, 1987.

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

[DeMe90] J. Desel and A. Merceron. P/T-systems as abstractions of C/E-systems. G. Rozenberg (ed.): Advances in Petri Nets 1989, Vol. 424 of Lecture Notes in Computer Science, Springer-Verlag, 1989.

[Dese98] J. Desel. Basic linear algebraic techniques for place/transition nets. Advanced Course on Petri Nets, 1996, in this volume.

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

[Espa98] J. Esparza. Decdability and complexity of Petri net problems – an introduction. Advanced Course on Petri Nets, 1996, in this volume.

[Golt90] U. Goltz. On condition/event representation of place/transition nets. K. Voss, H.J. Genrich and P.S. Thiagarajan (eds.): Concurrency and Nets, pp. 217–231, Springer-Verlag, 1990.

[GoRe83] U. Goltz and W. Reisig. The non-sequential behaviour of Petri nets. Information and Control, 3:125–147, 1983.

[Jant87] M. Jantzen. Complexity of place/transition nets. In [BRR87], pp. 413–435.

[JaVa80] M. Jantzen and R. Valk. Formal properties of place/transition nets. In [Brau80], pp. 165–21.

[Jens98] K. Jensen. An introduction to the practical use of coloured Petri nets. Advanced Course on Petri Nets, 1996, in this volume.

[Kosa82] S.R. Kosaraju. Decidability of reachablity in vector addition systems. 14th Annual ACM Symposium on Theory of Computing, San Francisco, pp. 267–281, 1982.

[Laut87] K. Lautenbach. Linear algebraic techniques for place/transition nets. In [BRR87], pp. 142-167.

[Mayr84] E.W. Mayr. An algorithm for the general Petri net reachability problem. SIAM Journal of Computing, 13(3):441-459, 1984.

[Mura89] T. Murata. Petri net: properties, analysis and application. Proceedings of the IEEE 77(4):541–580, 1989.

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

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

[Reis87] W. Reisig. Place/transition systems. In [BRR87], pp. 117-141.

[Reut88] C. Reutenauer. The Mathematics of Petri Nets. Masson (1988).

[RoEn98] G. Rozenberg and J. Engelfriet. Elementary net systems. Advanced Course on Petri Nets, 1996, in this volume.

[Valm98] A, Valmari. The state explosion problem. Advanced Course on Petri Nets, 1996, in this volume.