Characterization of Petri Net Languages

Jörg P. Bachmann

Humboldt-Universität zu Berlin

Abstract. This work treats the relationship between Petri Nets and languages. In 2003, Darondeau solved the Petri Net synthesis problem for a certain class of languages, i. e. for a given language $\mathcal{L}$ of this class he constructed a Petri Net generating the least Petri Net language containing $\mathcal{L}$. Furthermore, he developed a decision procedure which tests for equality of these two languages. It is proved here that all Petri Net languages are in the class of languages introduced by Darondeau. Namely, these are the prefix-closed languages which have a semi-linear commutative (also called Parikh) image. Hence, a characterization of Petri Net languages is given.

Keywords: Petri Net, P/T-Net, language, characterization, synthesis

1 Introduction

A Petri Net generates a language in a natural way. Considering the set of transitions as an alphabet, the language generated by a given Petri Net is the set of all feasible transition sequences in that net. Given a language, the question is raised whether a generating Petri Net exists and how it could possibly be constructed. This exactly is the Petri Net synthesis problem for languages. In [1], P. Darondeau solved this problem for a class of languages with certain properties. This work shows that all languages, generated by a Petri Net, fulfill these properties, i. e. no other languages with a generating Petri Net exist.

1.1 Notation

Let $\mathbb{N} = \{0, 1, 2, \ldots\}$ denote the set of nonnegative integers. $\mathbb{Z} = (-\mathbb{N}) \cup \mathbb{N}$ is the set of integers. $\mathbb{Q} = \left\{ \frac{p}{q} \mid p, q \in \mathbb{N}, q \neq 0 \right\}$ is the set of rational numbers. Let $n \in \mathbb{N} \setminus \{0\}$ be a positive integer. $\mathbb{Z}/n\mathbb{Z}$ is the factor ring with n elements. $\mathbb{N}^n$ is the Cartesian product of $\mathbb{N}$ with itself n times. For $v \in \mathbb{N}^n$ and $1 \leq i \leq n$ the i-th component of v is denoted by v_i. The Kronecker delta δ_{ij} is 1 iff i equals j and 0 otherwise. a divides b (written by $a \big| b$) holds iff $\exists c \in \mathbb{N} : b = a \cdot c$.

If Σ is an alphabet, then Σ^* is the set of all words of letters in Σ. A language $\mathcal{L}$ over Σ is a subset of Σ^*. ε denotes the empty word. Especially, for a semi-group $\mathcal{M}$ with binary operation '$\cdot$' and a subset $F \subseteq \mathcal{M}$, let $F^* := \{a_1 \cdot \ldots \cdot a_k \mid k \in \mathbb{N}, a_1, \ldots, a_k \in \mathcal{M}\}$.

Let $f : A \longrightarrow B$ be a mapping and $A' \subseteq A$ be a subset. Then, $f\big|_{A'} : A' \longrightarrow B$ with $a \longmapsto f(a)$ is the restriction of f to A'.

Let $a, b \in \mathbb{N}$ be two nonnegative integers, such that $a \leq b$. Then, $[a, b] := \{a, \ldots, b\} \subseteq \mathbb{N}$ is the interval from a to b including a and b.

CONCURRENCY, SPECIFICATION AND PROGRAMMING
M. Szczuka et al. (eds.): Proceedings of the international workshop CS&P 2011
September 28-30, Pułtusk, Poland, pp. 17-28

1.2 Basic Definitions and Facts

Definition 1 (Petri Net). *A marked Petri Net is a finite, directed, weighted and two-colored graph* $\mathcal{N} = (P, T, F, m)$ *where* P *is the set of places,* T *is the set of transitions,* $F : (P \times T) \cup (T \times P) \longrightarrow \mathbb{N}$ *represents the weights,* $M : P \longrightarrow \mathbb{N}$ *is the initial marking and the sets* P *and* T *are finite and disjunct. For a place* $p \in P$*,* $m(p)$ *is called the number of tokens on* p*.*

The initial marking of a Petri Net will mostly be called m_0. From now on, $\mathcal{N} = (P, T, F, m_0)$ is a marked Petri Net.

Definition 2 (prespace, preplaces, precondition). *Let* $t \in T$ *be an arbitrary transition. The set* ${}^\bullet t := \{p \in P \mid F(p, t) > 0\}$ *is called prespace of* t *whereas its elements are called preplaces of* t*. The set* $\{F(p, t) \mid p \in {}^\bullet t\}$ *is called precondition of* t*. If* $P = \{p\}$ *consists of only one element, then the precondition is* $F(p, t)$ *for short.*

Definition 3 (fire, feasible transition sequence). *Let* $t \in T$ *be an arbitrary transition.* t *is concessionary (in the given marking* m*) iff* $\forall p \in {}^\bullet t : m(p) \geq F(p, t)$*.*

If t *is concessionary, then the firing of* t *means the transformation of the marking* m *of the given Petri Net* $\mathcal{N}$ *into a new marking* m'*, where* $M'(p) = M(p) + F(t, p) - F(p, t)$*. This is denoted by* $m \xrightarrow{\;t\;} m'$*.*

A transition sequence $v = t_1 \cdot \ldots \cdot t_k$ *is a concatenation of transitions* $t_i \in T$*.* v *is feasible iff* $m_0 \xrightarrow{t_1} m_1 \xrightarrow{t_2} \ldots \xrightarrow{t_n} m_n$ *and each* t_i *is concessionary in* m_{i-1}*.*

Definition 4. $\mathcal{L}(\mathcal{N})$ *is the language of all feasible transition sequences in* $\mathcal{N}$*.*

Definition 5 (subnet, atomic Petri Net). *A Petri Net* $\mathcal{N}' = (P', T', F', M')$ *is a subnet of* $\mathcal{N} = (P, T, F, M)$*, denoted by* $\mathcal{N}' \subseteq \mathcal{N}$*, iff the set of places is a subset and all other sets are the appropriate restrictions, i. e.* $P' \subseteq P$*,* $T' = T$*,* $F' = F\big|_{(P' \times T') \cup (T' \times P')}$ *and* $M' = M\big|_{P'}$*. Atomic Petri Nets are Petri Nets with exactly one place.*

Proposition 1. $\mathcal{L}(\mathcal{N})$ *is the intersection of the languages of all (atomic) subnets of* $\mathcal{N}$*, i. e.* $\mathcal{L}(\mathcal{N}) = \bigcap_{\mathcal{N}' \subseteq \mathcal{N}} \mathcal{L}(\mathcal{N}')$*.*

Proof. $\mathcal{L}(\mathcal{N}) \subseteq \bigcap_{\mathcal{N}' \subseteq \mathcal{N}} \mathcal{L}(\mathcal{N}')$ is obvious since $\mathcal{L}(\mathcal{N}) \subseteq \mathcal{L}(\mathcal{N}')$ holds for each (atomic) subnet $\mathcal{N}'$ of $\mathcal{N}$.

The proof of the other direction will be done by induction on the length of $w \in \bigcap_{\mathcal{N}' \subseteq \mathcal{N}} \mathcal{L}(\mathcal{N}')$.

Basis: Since $\varepsilon \in \mathcal{L}(\mathcal{N})$ holds for each Petri Net $\mathcal{N}$, the case of $w = \varepsilon$ is obvious.

Inductive step: Let $w \cdot t \in \bigcap_{\mathcal{N}' \subseteq \mathcal{N}} \mathcal{L}(\mathcal{N}')$. $w \in \mathcal{L}(\mathcal{N})$ holds by induction hypothesis. Consider the preplaces of t in $\mathcal{N}$.

- If ${}^\bullet t = \emptyset$, then t is always able to fire in $\mathcal{N}$, thus $w \cdot t \in \mathcal{L}(\mathcal{N})$.
- If, on the other hand, ${}^\bullet t \neq \emptyset$, then choose $p \in {}^\bullet t$ arbitrarily and consider $\mathcal{N}' = (\{p, \ldots\}, T, F', m_0') \subseteq \mathcal{N}$. Let $\hat{m}$ be the marking, such that $m_0 \xrightarrow{w} \hat{m}$ in $\mathcal{N}$. Furthermore, let m' be the marking, such that $m_0' \xrightarrow{w} m'$ in $\mathcal{N}'$. Then $m'(p) = \hat{m}(p)$ holds. $w \cdot t \in \mathcal{L}(\mathcal{N}')$ implies $\hat{m}(p) = m'(p) \geq F(p, t)$. Since $p \in {}^\bullet t$ has been chosen arbitrarily, t is concessionary in $\mathcal{N}$ after w fired. Hence, $w \cdot t \in \mathcal{L}(\mathcal{N})$ holds. Furthermore, $w \cdot t \in \bigcap \mathcal{L}(\mathcal{N}')$ has been chosen arbitrarily, thus $\bigcap_{\mathcal{N}' \subseteq \mathcal{N}} \mathcal{L}(\mathcal{N}') \subseteq \mathcal{L}(\mathcal{N})$ holds. □

2 Characterization of Petri Net Languages

2.1 Atomic Petri Nets

Definition 6 (producing, consuming, value). *Let $\mathcal{N} = (\{p\}, T, F, m_0)$ be an atomic Petri Net.*

- *A transition $t \in T$ is called producing iff $F(p, t) \leq F(t, p)$. Otherwise t is called consuming.*
- *Let $u \in T^*$ be a transition sequence. The value of $u = t_1 \cdot \cdots \cdot t_\ell$ is*

$$V(u) := \sum_{i=1}^{\ell} \Big(F(t_i, p) - F(p, t_i) \Big) .$$

  Especially, the value of a transition $t \in T$ is $V(t) = F(t, p) - F(p, t)$.
- *The value of a Parikh vector v is the value of a transition sequence u with $[u] = v$.*

It's easy to see that the last point of definition 6 is well defined, i. e. for two transition sequences $u \in T^*$ and $u' \in T^*$ with $[u] = [u']$ holds $V(u) = V(u')$. If $v = [u]$, then

$$V(u) = \sum_{i=1}^{n} v_i \cdot \Big(F(t_i, p) - F(p, t_i) \Big) .$$

Hence, for a given Petri Net $\mathcal{N}$, the mapping $V : T^* \longrightarrow \mathbb{N}^n$ is a homomorphism, i. e. $V(u \cdot u') = V(u) + V(u')$ for $u, u' \in T^*$.

Definition 7 ($\mathcal{N}^T$). *Let $\mathcal{N} = (P, T, F, m_0)$ be a Petri Net and $T' \subseteq T$ be a subset of transitions. $\mathcal{N}^{T'}$ is the Petri Net which arises from $\mathcal{N}$ by restricting the set of transitions to T', i. e.*

$$\mathcal{N}^{T'} = \Big\{ P, T', F|_{(P \times T') \cup (T' \times P)}, m_0 \Big\} .$$

Definition 8 (normal form). *Let $\mathcal{N}$ be an atomic Petri Net and $v = t_1 \cdot \ldots \cdot t_n \in \mathcal{L}(\mathcal{N})$ be a transition sequence. v is in normal form iff the following three properties hold:*

- *The transition sequence can be split into $v = p \cdot s$, such that p consists of only producing transitions and s consists of only consuming transitions.*
- *If $p = t_1 \cdot \ldots \cdot t_m$ is the prefix, then $i < j$ for $i, k \in [1, m]$ implies $F(p, t_i) \leq F(p, t_j)$, i. e. the transitions are ordered by ascending precondition.*
- *If $s = t_{m+1} \cdot \cdots \cdot t_n$ is the suffix, then $i < j$ for $i, j \in [m + 1, n]$ implies $F(p, t_i) \geq F(p, t_j)$, i. e. the transitions are ordered by descending precondition.*

Proposition 2. *Let $\mathcal{N} = \{\{p\}, T, F, m_0\}$ be an atomic Petri Net and $u \in [\mathcal{L}(\mathcal{N})]$ be a Parikh vector in the Parikh image of $\mathcal{L}(\mathcal{N})$. Then, there is a feasible transition sequence $v \in \mathcal{L}(\mathcal{N})$ in normal form, such that $[v] = u$.*

Proof. Let $T = \{t_1, \ldots, t_n\}$ denote the set of transitions in $\mathcal{N}$. Let $v \in \mathcal{L}(\mathcal{N})$ be an arbitrary feasible transition sequence with $[v] = u$. The proof will be done by induction on the length $k := |v|$.

Basis ($k = 0$): $v = \varepsilon$ is the only feasible transition sequence with $|v| = 0$. Trivially, this transition sequence is in normal form.

Inductive step ($k \to k + 1$): Let

$$T' := \{t_j \in T \mid [v]_j \neq 0\}$$

be the set of transitions in T which occur in v. A case differentiation on the type of transitions in T' will be made.

- If $\forall t \in T' : F(p, t) \geq 0$ holds, then choose $m \in \mathbb{N}$ with $t_m \in T'$ and

  $$F(p, t_m) = \max_{t \in T'} F(p, t) .$$

  Then, $v = b \cdot t_m \cdot s$ can be split, such that t_m does not occur in b. Because of the maximality of $F(p, t_m)$ in the set of transitions $t \in T'$, all transitions are concessionary after b fired, i. e.

  $$\forall t \in T' : F(p, t) \leq m_0 + V(b) .$$

  Since all transitions $t \in T'$ are producing, the unequation

  $$F(p, t_m) \leq m_0 + V(b) \leq m_0 + V(b) + V(s) = m_0 + V(b \cdot s)$$

  holds. Moreover, $b \cdot s \in \mathcal{L}(\mathcal{N})$ is a feasible transition sequence. By induction hypothesis, there is a $w \in \mathcal{L}(\mathcal{N})$ in normal form, such that $[b \cdot s] = [w]$. It follows

  $$F(p, t_m) \leq m_0 + V(b \cdot s) = m_0 + V(w) ,$$

  i. e. t_m is also concessionary after w fired. Hence, $w \cdot t_m \in \mathcal{L}(\mathcal{N})$ is a feasible transition sequence. In the end, $w \cdot t_m$ is in normal form and $[w \cdot t_m] = [v]$ holds as wished.

– If, on the other hand, $\exists t \in T' : F(p,t) < 0$ holds, then choose $m \in \mathbb{N}$ with $t_m \in T'$ and

$$F(p, t_m) = \min_{t \in T'} F(p, t)\,.$$

Split $v = b \cdot t_m \cdot s$, such that t_m does not occur in b. Since t_m is a consuming transition, the unequation

$$m_0 + V(b) > m_0 + V(b \cdot t_m)$$

holds. Hence, $b \cdot s \in \mathcal{L}(\mathcal{N})$ is a feasible transition sequence. Let t_ℓ be the last transition of s and s' be the prefix of s, such that $b \cdot s = b \cdot s' \cdot t_\ell$. $F(p,t_m) \leq F(p,t_\ell)$ holds, based on the choice of $t_m \in T'$. Since $v = p{\cdot}t_m{\cdot}s'{\cdot}t_\ell$ is a feasible transition sequence, the following unequation holds:

$$\begin{aligned}
& & M_0 + V(p \cdot t_m \cdot s') &\geq F(p,t_\ell) \\
&\Longrightarrow & M_0 + V(p \cdot s') &\geq F(p,t_\ell) - V(t_m) \\
&\Longrightarrow & M_0 + V(p \cdot s' \cdot t_\ell) &= M_0 + V(p \cdot s') + V(t_\ell) \\
& & &\geq F(p,t_\ell) - V(t_m) + V(t_\ell) \\
& & &= F(p,t_m) + \underbrace{F(t_\ell,p) - F(t_m,p)}_{\geq 0} \\
& & &\geq F(p,t_m)\,.
\end{aligned}$$

By induction hypothesis, there is a $w \in \mathcal{L}(\mathcal{N})$ in normal form, such that $[p \cdot s] = [p \cdot s' \cdot t_\ell] = [w]$. Hence,

$$M_0 + V(w) = M_0 + V(p \cdot s) = M_0 + V(p \cdot s' \cdot t_\ell) \geq F(p,t_m)\,,$$

i. e. t_m is concessionary after w fired, thus $w \cdot t_m \in \mathcal{L}(\mathcal{N})$ is a feasible transition sequence too. In the end, $w \cdot t_m$ is in normal form, because of the choice of t_m and $[w \cdot t_m] = [v]$ holds as wished. □

2.2 Semi-linearity of Parikh Images of Petri Net Languages

Lemma 1. *Let $M := \{a_1, \ldots, a_n\} \subset \mathbb{N} \setminus \{0\}$ be a finite set of positive integers. Then, there is a subset $T = \{a_{i_1}, \ldots, a_{i_m}\} \subseteq M$, such that*

$$n \,\Big|\, \sum_{j=1}^{m} a_{i_j}\,. \tag{1}$$

Proof. Consider the function

$$\begin{aligned}
f : [1,n] &\longrightarrow \mathbb{Z}/n\mathbb{Z} \\
r &\longmapsto \sum_{j=1}^{r} a_j \mod n\,.
\end{aligned}$$

If f is injective, then there is an $r \in [1,n]$ with $f(r) = 0 \mod n$. This implies $n \big| \sum_{j=1}^{r} a_j$ and $T := \{a_1, \ldots, a_r\}$ fulfills (1).

Otherwise, there are numbers $r, s \in [1,n]$ with $r < s$ and $f(r) = f(s)$. Then, $f(s) - f(r) = 0 \mod n$ and thereby $n \big| \sum_{j=r+1}^{s} a_j$ holds. Hence, $T := \{a_{r+1}, \ldots, a_s\}$ fulfills (1). □

Lemma 2. *The Parikh image of each language, generated by a Petri Net, is semi-linear.*

Proof. [3] shows that the intersection of semi-linear languages is semi-linear again. Because of that and Proposition 1, it suffices to prove that the Parikh image of the language of an arbitrary atomic Petri Net is semi-linear. Therefor, let

$$\mathcal{N} = (\{p\}, T, F, m_0)$$

be an arbitrary atomic Petri Net. The proof of the semi-linearity of the Parikh image of $\mathcal{N}$ will be shown by induction on the number $n := |T|$ of transitions in $\mathcal{N}$.

Basis ($n = 0$): If $T = \emptyset$, then $\mathcal{L}(\mathcal{N}) = \emptyset$ holds, whereby $[\mathcal{L}(\mathcal{N})]$ is semi-linear.

Inductive step ($n - 1 \longrightarrow n$): Denote $T = \{t_1, \ldots, t_n\}$ with the following property: If there is a consuming transition $t \in T$, then t_n is such a transition with minimal precondition $F(p,t)$. Otherwise, t_n is an arbitrary transition with maximum precondition $F(p,t)$. By induction hypothesis, the Parikh image of $\mathcal{N}^{T'}$ is semi-linear for $T' = \{t_1, \ldots, t_{n-1}\}$. Denote

$$S' := \left[\mathcal{L}\left(\mathcal{N}^{T'}\right)\right] = \bigcup_{i=1}^{k} e_i \cdot F_i^* .$$

A case differentiation on the type of t_n will be made.

- Consider t_n to be a producing transition. By construction of the case differentiation, this implies that all transitions $t_i \in T$ are producing.
  If t_i is already concessionary in the initial marking, then all transitions $t_i \in T$ are concessionary in the initial marking, since t_n had maximum precondition. Thus, every transition is able to fire in the initial and each reachable marking, whereby

$$\left[\mathcal{L}\left(\mathcal{N}^{T}\right)\right] = \Big\{[t_1], \ldots, [t_n]\Big\}^* \tag{2}$$

  holds.
  If, on the other hand, t_i is not concessionary in the initial marking, then S' has to be extended by some more linear sets. Note, that S' consists of vectors in $\mathbb{N}^{n-1}$. If necessary, these are embedded trivially in $\mathbb{N}^n$ by setting

the additional component to zero. Consider an arbitrary linear set $e_i \cdot F_i^*$ of S' and set

$$F_i' := \{f \in F_i \mid V(f) > 0\} \subseteq F_i \,.$$

Since $V(f) > 0$ holds for each $f \in F_i'$, there are only a finite number of feasible transition sequences in $\mathcal{N}$, whose Parikh images are in $e_i \cdot (F_i')^*$ and which make t_n concessionary for the first time (i. e. no prefix makes t_n concessionary). Let W_i be the set of Parikh images of these transition sequences. The sets

$$S_i := \bigcup_{w \in W_i} w \cdot \Big(F_i \cup \big\{[t_n]\big\}\Big)^* \tag{3}$$

extend S' to S, i. e.

$$S := S' \cup \left(\bigcup_{i=1,\ldots,k} S_i \right) \,.$$

Proof of correctness: It has to be proved yet that this construction is correct in the sense of $S = \left[\mathcal{L}\left(\mathcal{N}^T\right)\right]$. $S \subseteq \left[\mathcal{L}\left(\mathcal{N}^T\right)\right]$ holds by construction. Note that all transitions $t \in T$ are concessionary after firing of feasible transition sequences with Parikh vectors $w \in W_i$, because $t_n \in T$ has maximum precondition.
It remains to show that $\left[\mathcal{L}(\mathcal{N}^T)\right] \subseteq S$ holds. Therefor, let $v \in \mathcal{L}\left(\mathcal{N}^T\right)$ be an arbitrary feasible transition sequence in $\mathcal{N}^T$. If t_n does not occur in v, then $v \in \mathcal{L}(\mathcal{N}^{T'})$ holds and thereby $[v] \in S' \subseteq S$. Otherwise, there is a shortest prefix q of v which makes t_n concessionary. Let q' be the feasible transition sequence which arises from q by removing all transitions t_j with $V(t_j) = 0$. Furthermore, let b be a transition sequence, such that $[q] = [q'] + [b]$, i. e. b consists of all of the from q removed transitions. It follows that $V(q) = V(q')$ and q' makes t_n concessionary for the first time too. Hence, there is an i with $1 \leq i \leq k$, such that $[q'] \in W_i$. Then,

$$[v] \in [q'] \cdot \Big(F_i \cup \big\{[t_n]\big\}\Big)^* \subset S$$

holds, because $[b] \in F_i^*$.[1] Since $v \in \mathcal{L}\left(\mathcal{N}^T\right)$ has been chosen arbitrarily, $\left[\mathcal{L}\left(\mathcal{N}^T\right)\right] \subseteq S$ and hence the equality holds.

– Consider t_n to be a consuming transition. Let $e_i \cdot F_i^*$ be an arbitrary linear set of S' and set

$$G_i := \left\{ g = \sum_{j=1}^{r} f_j + m \cdot [t_n] \mid r \leq |V(t_n)|, f_j \in F_i, m \in \mathbb{N}, V(g) \geq 0, \right\}.$$

[1] Cf. (2) and (3).

G_i contains F_i, since $|V(t_n)| \geq 1$ and $V(f) \geq 0$ holds for $f \in F_i$. r is limited by $|V(t_n)|$. Since t_n is a consuming transition and $V(g) \geq 0$ holds for $g \in G_i$, m is limited by $|V(t_n)|$ too. Hence, the finiteness of F_i implies the finiteness of G_i.

The condition $V(g) \geq 0$ benefits that parts of transition sequences with Parikh vector in G_i can be fired any number of times, as long as they can be fired once.

Possibly, $V(e_i) > 0$ and $m_0 > 0$ holds and there is a feasible transition sequence u, such that $m_0 + V(u) \geq F(p, t_n)$, and $[u] - e_i \notin G_i$. Then $u \cdot t_n \in \mathcal{L}(\mathcal{N}^T)$ would hold, but nevertheless $[u \cdot t_n] \notin e_i \cdot G_i^*$. So the initial marking and the value of e_i have to be observed too. Please note that possibly t_n cannot fire until there are no tokens left on the place p, since $F(t_n, p) > 0$ might hold. Thus, the affine vector e_i of $e_i \cdot F_i^*$ has to be combined with $[t_n]$ appropriate. Therefor, let

$$\begin{aligned} H_i := \Bigg\{ h = e_i + \sum_{j=1}^{r} f_j + m \cdot [t_n] \mid r < 2 \cdot F(p, t_n), f_j \in F_i, \\ m_0 + V\left(e_i + \sum_{j=1}^{r} f_j\right) \geq F(p, t_n), \\ m \in \mathbb{N}, m \neq 0 \rightarrow m_0 + V(h) - V(t_n) \geq F(p, t_n) \Bigg\} . \end{aligned} \tag{4}$$

Because of the limitedness of r, the finiteness of F_i and the last condition in (4), H_i is finite for consuming transitions t_n. The last condition of (4) regards that t_n is able to fire only if there are enough tokens on the place p. Now, every case is conceived. Thereby, let

$$S_i := \bigcup_{h \in H_i} h \cdot G_i^*$$

and finally

$$S := S' \cup \left(\bigcup_{i=1,\ldots,k} S_i \right) .$$

Proof of correctness: Again, it has to be proved that this construction is correct in the sense of $S = \left[\mathcal{L}\left(\mathcal{N}^T\right)\right]$. It remains to show that $\left[\mathcal{L}(\mathcal{N}^T)\right] \subseteq S$ holds, since the other direction holds by construction. Therefor, let $v \in \mathcal{L}(\mathcal{N}^T)$ be an arbitrary feasible run and v' be the sequence, which arises from v by removing all occurrences of t_n. Since t_n is consuming, $v' \in \mathcal{L}(\mathcal{N}^{T'})$ holds. By induction hypothesis, $[v'] \in e_i \cdot F_i^* \subseteq S'$ holds for a certain i. Let such an i be fixed. Then,

$$[v'] = e_i + f_1 + \cdots + f_r + f_{r+1} + \cdots + f_{r'}$$

holds for certain $f_1, \dots, f_{r'} \in F_i$ and there is a $c \in \mathbb{N}$, such that $[v] = [v'] + [t_n^c]$. Denote these f_j in such a way that $V(f_j) > 0$ holds for $j \leq r$ and $V(f_j) = 0$ holds for $j > r$.
By induction on r, it will be shown that $[v] = [v'] + [t_n^c] \in h \cdot G_i^* \subseteq S_i \subseteq S$ holds for a certain $h \in H_i$.

Basis: Let $r < 2 \cdot F(p, t_n)$. Then, set

$$u := [v] - \sum_{j=r+1}^{r'} f_j = [v'] - \sum_{j=r+1}^{r'} f_j + c \cdot [t_n] \quad = \left(e_i + \sum_{j=1}^{r} f_j \right) + c \cdot [t_n] \,.$$

If $c = 0$, then $u \in \left[\mathcal{L}\left(\mathcal{N}^{t'}\right)\right]$ and thereby $u \in S' \subseteq S$ holds.
Otherwise, $m_0 + V(e_i + \sum_{j=1}^{r} f_j) \geq F(p, t_n)$ holds, since t_n couldn't fire else.
The definition of H_i and $v \in \mathcal{L}(\mathcal{N}^T)$ imply $u \in H_i$.
Finally, $[v] \in h \cdot G_i^* \subseteq S_i \subseteq S$ holds for a certain $h \in H_i$, because of $\sum_{j=r+1}^{r'} f_j \in F_i^* \subseteq G_i^*$.

Inductive step: Consider $r \geq 2 \cdot F(p, t_n)$. By lemma 1, there are pairwise distinct $j_1, \dots, j_s \in [r - |V(t_n)| + 1, r]$ with

$$\sum_{\ell=1}^{s} V\left(f_{j_\ell}\right) = m \cdot |V(t_n)|$$

for certain $s \leq |V(t_n)|$ and $m \in \mathbb{N}$.
If $m \geq c$, then $[v]$ can be split into

$$\begin{aligned} [v] &= [v'] + c \cdot [t_n] \\ &= \underbrace{e_i + \sum_{\substack{j \in [1, r'] \\ j \notin \{j_1, \dots, j_s\}}} f_j}_{=:u} + \underbrace{\sum_{\ell=1}^{s} f_{j_\ell} + c \cdot [t_n]}_{=:u'} \,. \end{aligned}$$

$u' \in G_i$ holds by definition of G_i. By induction hypothesis, $u \in h \cdot G_i^*$ holds for a certain $h \in H_i$ and thereby $[v] \in h \cdot G_i^* \subseteq S_i \subseteq S$.
If, on the other hand, $m < c$, then $[v]$ can be split into

$$\begin{aligned} [v] &= [v'] + c \cdot [t_n] \\ &= e_i + \sum_{j=1}^{r'} f_j + c \cdot [t_n] \\ &= \underbrace{e_i + \overbrace{\left(\sum_{\substack{j \in [1, r'] \\ j \notin \{j_1, \dots, j_s\}}} f_j \right)}^{=:w} + (c - m) \cdot [t_n]}_{=:u} + \underbrace{\overbrace{\left(\sum_{\ell=1}^{s} f_{j_\ell} \right)}^{w'} + m \cdot [t_n]}_{=:u'} \end{aligned}$$

and $V(u') = 0$ holds. Assume that $u \in \left[\mathcal{L}(\mathcal{N}^T)\right]$. Then, $u \in h \cdot G_i^*$ holds for a certain $h \in H_i$ by induction hypothesis and $u' \in G_i$ holds by definition of G_i. That is, why $[v] = u + u' \in h \cdot G_i^* \subseteq S_i \subseteq S$.
It remains to show that $u \in \left[\mathcal{L}\left(\mathcal{N}^T\right)\right]$. By proposition 2, there is a $\bar{v} \in \mathcal{L}(\mathcal{N}^T)$ in normal form, such that $[\bar{v}] = [v]$. Since $\bar{v}$ is in normal form, t_n only appears at the end of the sequence w.l.o.g., i. e.

$$\exists \hat{v} \in \mathcal{L}(\mathcal{N}^{T'}) : \bar{v} = \hat{v} \cdot t_n^c \,.$$

By induction hypothesis,

$$e_i + w \in e_i \cdot F_i^* \subseteq \left[\mathcal{L}\left(\mathcal{N}^{T'}\right)\right]$$

holds. Again, by proposition 2, there is a $v^* \in \mathcal{L}(\mathcal{N}^{T'})$ in normal form, such that

$$[v^*] = e_i + w \,.$$

Then, $[v^* \cdot t_n^{c-m}] = e_i + w + (c-m) \cdot [t_n] = u$. Hence, it remains to show that $v^* \cdot t_n^{c-m} \in \mathcal{L}(\mathcal{N}^T)$. Since $v \in \mathcal{L}(\mathcal{N}^T)$ and t_n fires at last in $\bar{v}$,

$$V(v) = V(\bar{v}) \geq F(t_n, p)$$

holds, thus for $0 \leq \ell < c - m$ follows

$$\begin{aligned} V(v^* \cdot t_n^\ell) &= V([v] - (c-\ell) \cdot [t_n] - w') \\ &= \underbrace{V(v)}_{\geq F(t_n,p)} - \underbrace{V(w' + m \cdot [t_n])}_{=V(u')=0} - V\left((c-m-\ell) \cdot [t_n]\right) \\ &\geq F(t_n, p) - (c-m-\ell) \cdot V(t_n) \,. \end{aligned}$$

Because of $V(t_n) < 0$, the last term takes its minimum at $\ell = c - m - 1$, i. e.

$$\begin{aligned} V(v^* \cdot t_n^\ell) &\geq F(t_n, p) - (c - m - (c - m - 1)) \cdot V(t_n) \\ &= F(t_n, p) - V(t_n) \\ &= F(p, t_n) \,. \end{aligned}$$

This means that t_n is concessionary after firing each $v^* \cdot t_n^\ell \in \mathcal{L}(\mathcal{N}^T)$ for $0 \leq \ell < c - m$. Thereby, $v^* \cdot t_n^{c-m} \in \mathcal{L}(\mathcal{N}^T)$ and $[v^* \cdot t_n^{c-m}] = u \in \left[\mathcal{L}(\mathcal{N}^T)\right]$ hold, as claimed.
Since $v \in \mathcal{L}(\mathcal{N}^T)$ had been chosen arbitrarily, $[\mathcal{L}(\mathcal{N}^T)] \subseteq S$ and hence the equality holds. □

2.3 Characterization of Petri Net Synthesis Languages

Lemma 3. *Let $e \cdot F^* \subseteq \mathbb{N}^n$ be an arbitrary linear set. Then, the restriction of the set $e \cdot (-\delta_{ji})_{1 \leq j \leq n} \cdot F^*$ to $\mathbb{N}^n$ is semi-linear.*

Proof. Set

$$S := e \cdot (-\delta_{ji})_{1 \leq j \leq n} \cdot F^* \cap \mathbb{N}^n \,.$$

Let $v_1, \ldots, v_k$ be the elements of F, which have a positive value in the i-th component. It will be revealed that

$$S = \bigcup_{j=1}^{k} e \cdot (-\delta_{ji})_{1 \leq j \leq n} \cdot v_j \cdot F^* =: T \,.$$

- $T \subseteq S$ is obvious since $v_j \in F$ and, by choice of $v_j \in F$, all vectors $(-\delta_{ji})_{1 \leq j \leq n} \cdot v_j$ have a positive entry in the i-th component.
- Choose $w \in S$ arbitrarily. Since $w \in \mathbb{N}^n$, the i-th component cannot be negative, i. e. $w_i \geq 0$. This only is possible if there is at least one $v_j \in F$ in the combination for w. That is, why $w \in e \cdot v_j \cdot (-\delta_{ji})_{1 \leq j \leq n} \cdot F^* \subseteq T$ holds for some j. Hence, $S \subseteq T$ holds. □

Corollary 1. *$[\mathcal{L}]$ is semi-linear iff $[\mathcal{L}/t]$ and $[\mathcal{L} \setminus (\mathcal{L}/t)]$ are semi-linear for each Transition t.*

Proof. Let $[\mathcal{L}] = \bigcup_{i=1}^{n} e_i \cdot F_i^*$ be semi-linear and $t \in T$. Then, by lemma 3

$$[\mathcal{L}/t] = \bigcup_{i=1}^{n} e_i \cdot (-\delta_{rj})_{1 \leq r \leq n} \cdot F_i^* \cap \mathbb{N}^n = \bigcup_{i=1}^{n} \bigcup_{s=1}^{k_i} e_i \cdot v_{i,s} \cdot F_i^*$$

is semi-linear too for some $v_{i,s} \in F$. By [3], the difference of semi-linear sets is semi-linear again. Thereby, $[\mathcal{L} \setminus (\mathcal{L}/t)] = [\mathcal{L}] \setminus [\mathcal{L}/t]$ is semi-linear.

On the other hand, $[\mathcal{L}] = [\mathcal{L}/t] \cup [\mathcal{L} \setminus (\mathcal{L}/t)]$ is semi-linear if $[\mathcal{L}/t]$ and $[\mathcal{L} \setminus (\mathcal{L}/t)]$ are semi-linear. □

Theorem 1. *Let $\mathcal{N} = (P, T, F, m_0)$ be an arbitrary Petri Net and $\mathcal{L} = \mathcal{L}(\mathcal{N})$ be the language, generated by $\mathcal{N}$. Then, $\mathcal{L}$ is prefix-closed and $[\mathcal{L}/t]$ and $[\mathcal{L} \setminus (\mathcal{L}/t)]$ are semi-linear for each Transition $t \in T$.*

Proof. Obviously, $\mathcal{L}$ is prefix-closed, since every prefix of a feasible transition sequence is feasible too.

Consider an arbitrary atomic subnet $\mathcal{N}'$ of $\mathcal{N}$. Set $\mathcal{L}' := \mathcal{L}(\mathcal{N}')$. Lemma 2 implies the semi-linearity of $[\mathcal{L}']$. Since $\mathcal{L} = \bigcap_{\mathcal{N}' \subseteq \mathcal{N}} \mathcal{L}(\mathcal{N}')$ is the intersection of the languages of all (atomic) subnets of $\mathcal{N}$ and by [3] the intersection of a finite number of semi-linear sets is semi-linear again,

$$[\mathcal{L}] = \left[\bigcap_{\mathcal{N}' \subseteq \mathcal{N}} \mathcal{L}(\mathcal{N}') \right] = \bigcap_{\mathcal{N}' \subseteq \mathcal{N}} [\mathcal{L}(\mathcal{N}')]$$

is semi-linear too. By corollary 1, $[\mathcal{L}/t]$ and $[\mathcal{L} \setminus (\mathcal{L}/t)]$ are semi-linear for each $t \in T$. □

Darondeaus decision procedure and construction of a Petri Net for a given language requires the conditions mentioned in Theorem 1. Theorem 1 shows that there is no Petri Net, generating a language, which does not fulfill at least one of these conditions.

Corollary 1 shows the equivalence of the semi-linearity of $[\mathcal{L}/t]$ and $[\mathcal{L} \setminus (\mathcal{L}/t)]$ for each Transition t to the semi-linearity of $[\mathcal{L}]$. Thereby, with [1] and this work, the class of Petri Net synthesis decidable languages (or short Petri Net languages) can be manifested.

Definition 9 (Petri Net languages). *Let*

$$\mathcal{PT} := \{\mathcal{L} \mid \mathcal{L} \textit{ is prefix-closed and } [\mathcal{L}] \textit{ is semi-linear}\}$$

be the set of Petri Net synthesis decidable languages.

3 Conclusion

This present work in combination with the work [1] of P. Darondeau completely solves the Petri Net synthesis problem for languages. The characterization given here determines languages for which it is decidable whether they have a generating Petri Net. But still, it might be possible to find an exact characterization of those languages having a generating Petri Net. It remains to analyze the complexity of algorithms finding such Petri Nets. Furthermore, the intersection of the class $\mathcal{PT}$, defined here, with other common classes like those in the Chomsky hierarchy could be studied. In [4], it has already been proved that context-free languages have a semi-linear Parikh image. In [2], the semi-linear language has even been constructed for a given context-free grammar. Hence, the class $\mathcal{PT}$ seems to be significant.

Moreover, the Petri Net synthesis problem for graphs has not completely been solved yet. A Petri Net generates its reachability graph. Thus, the Petri Net synthesis problem for graphs consists of finding and possibly constructing a Petri Net generating a given graph. Analogously to the Petri Net synthesis problem for languages, an exact characterization of the graphs having a generating Petri Net could possibly be given. Also, the complexity of algorithms constructing generating Petri Nets remains to be analyzed.

References

1. Philippe Darondeau, *Unbounded petri net synthesis*, Advanced Courses, 2003, pp. 413–438.
2. Javier Esparza, Pierre Ganty, Stefan Kiefer, and Michael Luttenberger, *Parikh's theorem: A simple and direct construction*, CoRR **abs/1006.3825** (2010).
3. Seymour Ginsburg and Edwin H. Spanier, *Bounded algol-like languages*, Transactions of The American Mathematical Society **113** (1964), 333–368.
4. Rohit J. Parikh, *On context-free languages*, J. ACM **13** (1966), 570–581.