

On Binary Words Being Petri Net Solvable

Kamila Barylska^{1,2*†}, Eike Best^{1*,**}, Evgeny Erofeev^{1***}, Lukasz Mikulski^{2†},
Marcin Piątkowski^{2†}

¹ Department of Comp. Sci., Carl von Ossietzky Univ. Oldenburg, Germany
`{eike.best,evgeny.erofeev}@informatik.uni-oldenburg.de`

² Faculty of Math. and Comp. Sci., Nicolaus Copernicus University Toruń, Poland
`{kamila.barylska,lukasz.mikulski,marcin.piatkowski}@mat.umk.pl`

Abstract. A finite word is called Petri net solvable if it is isomorphic to the reachability graph of some unlabelled Petri net. In this paper, the class of two-letter Petri net solvable words is studied.

Keywords: Binary words, labelled transition systems, Petri nets, synthesis

1 Introduction

Region theory [1] provides a polynomial algorithm, based on solving linear inequations, that checks whether a given finite labelled transition system is the reachability graph of some place/transition Petri net [5], and if it is, synthesises one of them. Due to the size of a transition system, this algorithm may be very time-consuming. Moreover, it may produce one out of a class of different nets, and there may not be a unique simplest one. For some applications, only certain types of labelled transition systems are relevant. This leads to the idea of investigating properties of labelled transition systems before synthesising them, in the hope of obtaining more efficient and possibly also more deterministic synthesis algorithms. It may even be possible to find exact structural characterisations, based solely on graph-theoretical properties, such as for the class of finite labelled transition systems which correspond to T-systems [2].

This paper reports progress on a similar effort about characterising the set of finite words over an alphabet $\{a, b\}$ which are Petri net solvable, i.e., for which a place/transition net with an isomorphic reachability graph exists. We shall put forward two conjectures, and describe some progress in analysing them. This

* Supported by a research and a visiting fellowship within the project "Enhancing Educational Potential of Nicolaus Copernicus University in the Disciplines of Mathematical and Natural Sciences" (project no. POKL.04.01.01-00-081/10)

** Supported by the German Research Foundation through the DFG/RFBR cooperation project CAVER, BE 1267/14-1

*** Supported by the German Research Foundation through the Research Training Group (DFG GRK 1765) SCARE – <http://www.uni-oldenburg.de/en/scare/>

† Supported by the National Science Center under grant no. 2013/09/D/ST6/03928

work could well be of interest in a wider context, as it might entail a nontrivial necessary condition for the solvability of an arbitrary labelled transition system. If the latter is solvable, then finding a PN-unsolvable word as a path in it may have a strong impact on its structure / shape.

2 Basic notations and conventions used in this paper

A *finite labelled transition system* with initial state is a tuple $TS = (S, \rightarrow, T, s_0)$ with nodes S (a finite set of states), edge labels T , edges $\rightarrow \subseteq (S \times T \times S)$, and an initial state $s_0 \in S$. A label t is enabled at $s \in S$, written formally as $s[t]$, if $\exists s' \in S: (s, t, s') \in \rightarrow$. A state s' is reachable from s through the execution of $\sigma \in T^*$, denoted by $s[\sigma]s'$, if there is a directed path from s to s' whose edges are labelled consecutively by σ . The set of states reachable from s is denoted by $[s]$. A (firing) sequence $\sigma \in T^*$ is allowed from a state s , denoted by $s[\sigma]$, if there is some state s' such that $s[\sigma]s'$. Two lts $TS_1 = (S_1, \rightarrow_1, T, s_{01})$ and $TS_2 = (S_2, \rightarrow_2, T, s_{02})$ are isomorphic if there is a bijection $\zeta: S_1 \rightarrow S_2$ with $\zeta(s_{01}) = s_{02}$ and $(s, t, s') \in \rightarrow_1 \Leftrightarrow (\zeta(s), t, \zeta(s')) \in \rightarrow_2$, for all $s, s' \in S_1$.

A *word over T* is a sequence $w \in T^*$, and it is *binary* if $|T| = 2$. A word $t_1 t_2 \dots t_n$ of length $n \in \mathbb{N}$ uniquely corresponds to a finite transition system $(\{0, \dots, n\}, \{(i-1, t_i, i) \mid 0 < i \leq n \wedge t_i \in T\}, T, 0)$.

An *initially marked Petri net* is denoted as $N = (P, T, F, M_0)$ where P is a finite set of places, T is a finite set of transitions, F is the flow function $F: ((P \times T) \cup (T \times P)) \rightarrow \mathbb{N}$ specifying the arc weights, and M_0 is the initial marking (where a marking is a mapping $M: P \rightarrow \mathbb{N}$, indicating the number of tokens in each place). A side-place is a place p with $p^\bullet \cap \bullet p \neq \emptyset$, where $p^\bullet = \{t \in T \mid F(p, t) > 0\}$ and $\bullet p = \{t \in T \mid F(t, p) > 0\}$. N is pure or side-place free if it has no side-places. A transition $t \in T$ is enabled at a marking M , denoted by $M[t]$, if $\forall p \in P: M(p) \geq F(p, t)$. The firing of t leads from M to M' , denoted by $M[t]M'$, if $M[t]$ and $M'(p) = M(p) - F(p, t) + F(t, p)$. This can be extended, as usual, to $M[\sigma]M'$ for sequences $\sigma \in T^*$, and $[M]$ denotes the set of markings reachable from M . The reachability graph $RG(N)$ of a bounded (such that the number of tokens in each place does not exceed a certain finite number) Petri net N is the labelled transition system with the set of vertices $[M_0]$, initial state M_0 , label set T , and set of edges $\{(M, t, M') \mid M, M' \in [M_0] \wedge M[t]M'\}$. If an lts TS is isomorphic to the reachability graph of a Petri net N , we say that N *solves* TS .

3 Separation problems, and an example

In region theory, a labelled transition system (S, \rightarrow, T, s_0) is assumed to be given as an input. In order to synthesise (if possible) a Petri net with isomorphic reachability graph, T is used as the set of transitions, and for the places, $\frac{1}{2} \cdot (|S| \cdot (|S| - 1))$ *state separation problems* and up to $|S| \cdot |T|$ *event/state separation problems* have

to be solved. A state separation problem consists of a set of states $\{s, s'\}$ with $s \neq s'$, and for each such set, one needs a place that distinguishes them. Such problems are always solvable for words; for instance, we might introduce a counting place which simply has j tokens in state j . An event/state separation problem consists of a pair $(s, t) \in S \times T$ with $\neg(s[t])$. For every such problem, we need a place p such that $M(p) < F(p, t)$ for the marking M corresponding to state s , where F refers to the arcs of the hoped-for net.

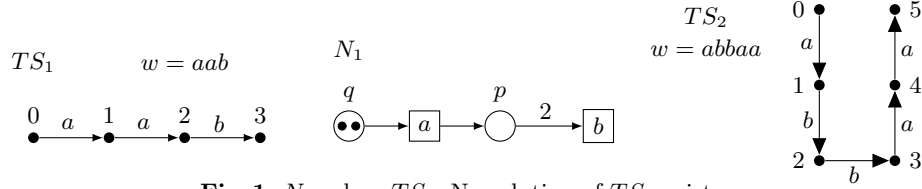


Fig. 1. N_1 solves TS_1 . No solution of TS_2 exists.

For example, in figure 1, the labelled transition systems TS_1 and TS_2 correspond to the words aab and $abbaa$, respectively. The former is PN-solvable, since the reachability graph of N_1 is isomorphic to TS_1 . TS_2 contains an unsolvable event/state separation problem. The state $s = 2$ just between the two b 's satisfies $\neg(s[a])$. We need a place q whose number of tokens in (the marking corresponding to) state 2 is less than needed for transition a to be enabled. Such a place q has the general form shown on the right-hand side of figure 2. It is useful to speak of the *effect* $\mathbb{E}(\tau)$ of a sequence $\tau \in T^*$ on place q . For the letter a , the effect is defined as $\mathbb{E}(a) = (a_+ - a_-)$, and this can be generalised easily. Thus, for instance, the effect $\mathbb{E}(abbaa)$ is $\mathbb{E}(abbaa) = 3 \cdot (a_+ - a_-) + 2 \cdot (b_+ - b_-)$. If q prevents a at state 2 in $abbaa$, then it must satisfy the following inequalities, amongst others: $a_- \leq m$, since a is enabled initially; $a_- \leq m + \mathbb{E}(abba)$, since state 4 enables a ; and $m + \mathbb{E}(ab) < a_-$, or equivalently, $0 \leq -m - \mathbb{E}(ab) + a_- - 1$, expressing the fact that q solves the event/state separation problem $\neg(2[a])$. Later, we show that this set of inequalities cannot be solved in the natural numbers.

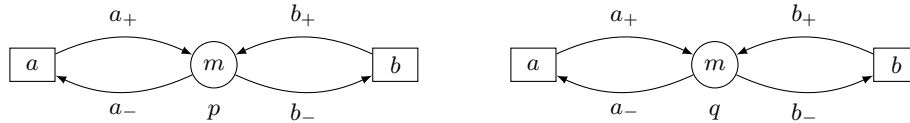


Fig. 2. Places p / q with four arc weights a_-, a_+, b_-, b_+ and initial marking m . They are similar, but p will be used for preventing b , and q for preventing a .

In a word of length n , the equation system for a single event/state separation problem comprises $n + 1$ inequations. In binary words, we have $n + 2$ such problems, one for every state $0, \dots, n - 1$ and two for the last state. Thus, a word w of length n is PN-solvable if and only if all those $n + 2$ systems, each having $n + 1$ inequations and five unknowns a_-, a_+, b_-, b_+, m , are solvable in \mathbb{N} . The question dealt with in this paper is whether the set of binary words that are PN-(un)solvable can be characterised equivalently, in a more structural way. We shall assume, from now on, that $T = \{a, b\}$.

4 Minimal unsolvable binary words, and some conjectures

Let a word $w' \in T^*$ be called a *subword* (or *factor*) of $w \in T^*$ if $\exists(u_1, u_2 \in T^*)$: $w = u_1 w' u_2$, and let $\#_t(w)$ denote the number of times the letter t occurs in w . Observe that if w is PN-solvable then all its subwords are, too. To see this, let the Petri net solving w be executed up to the state before w' , take this as the new initial marking, and add a pre-place with $\#_a(w')$ tokens to a and a pre-place with $\#_b(w')$ tokens to b . Thus, if a subword of w is unsolvable, then w is. For this reason, the notion of a *minimal unsolvable word* is well-defined (namely, as an unsolvable word all of whose subwords are solvable). A complete list of minimal unsolvable words up to length 110 can be found in [6]. As a consequence of the next proposition, any minimal unsolvable word either starts and ends with a or starts and ends with b .

Proposition 1. SOLVABILITY OF aw AND wb IMPLIES SOLVABILITY OF awb

If both aw and wb are solvable, then awb is also solvable.

Proof: Assume that aw and wb are PN-solvable words over $\{a, b\}$. If $w = b^k$ (for $k \in \mathbb{N}$) then $awb = ab^{k+1}$ is obviously solvable, hence we assume that b contains at least one a . Let $N_1 = (P_1, \{a, b\}, F_1, M_{01})$ and $N_2 = (P_2, \{a, b\}, F_2, M_{02})$ be Petri nets such that N_1 solves aw and N_2 solves wb . We can assume that N_1 and N_2 are disjoint, except for their transitions a and b . Forming the union of N_1 and N_2 gives a net which is synchronised at a and b , and which allows all (and only) sequences allowed by both N_1 and N_2 . We modify N_1 and N_2 before forming their union, as follows:

- (i) Modify N_1 by adding, to each place p in $\bullet b \cap P_1$, another $F_1(p, b)$ tokens. This allows an additional b .
- (ii) Modify N_2 by adding, to each place q in $\bullet a \cap P_2$, another $F_2(q, a)$ tokens. This allows an additional a . Further, for each place p in $a \bullet \cap P_2 \cap \bullet b$, add the quantity $F_2(a, p)$ both to $F_2(p, b)$ and to $F_2(b, p)$. The new arc weights lead to the same effect of b on p , but prevent premature occurrences of b which could have been allowed by adding the tokens in front of b in step (i).

Define N as the union of the two nets thus modified, and see figure 3 for an example. (The added tokens are drawn as hollow circles.) In general, N solves awb in the following way: The initial a is allowed in N_1 by definition and in N_2 by the additional tokens. The subsequent w is allowed in both nets, and hence in their synchronisation. The final b is allowed in N_2 by definition and in N_1 by the additional tokens. No premature b is allowed by the arc weight increase, and no final additional a is allowed because N_1 does not allow it. \square 1

From the list [6], it can be observed that all minimal unsolvable words starting and ending with a are of the following general form:

$$s_0 [(aba) b^*] s [(ba\alpha)^+] r [a] \quad \text{where } \alpha \in T^* \text{ and } s_0, s, r \text{ are states} \quad (1)$$

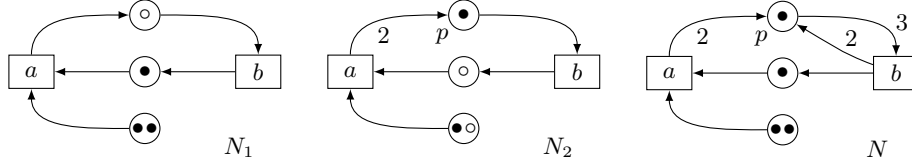


Fig. 3. N_1 (black tokens) solves $aw = abab$. N_2 (black tokens) solves $wb = babb$. N (redundant places omitted) solves $awb = ababb$.

with a not being separated at s . For example, $abbaa$ satisfies (1) with $\alpha = \varepsilon$, the star $*$ being repeated zero times, and the plus $+$ being repeated just once. Indeed, it is easy to prove that such words are generally PN-unsolvable:

Proposition 2. SUFFICIENT CONDITION FOR THE UNSOLVABILITY OF A WORD
If a word over $\{a, b\}$ has a subword of the form (1), then it is not PN-solvable.

Proof: For a word w of the form (1), we prove that w cannot be solved (implying the proposition in the context of the considerations above). Because $ba\alpha$ occurs at least once after state s , b is enabled at s , and a is not enabled at s . Suppose that some place q as in figure 2 (r.h.s.) exists which separates a at s . Let E be $\mathbb{E}(ab\alpha)$, i.e., the effect of $ab\alpha$ on q , and let $E_b = \mathbb{E}(b)$. $\mathbb{E}(b) \geq 1$ because q separates a at s but not at $s + 1$. For w , we derive the following inequalities:

$$\begin{aligned}
 (0) \quad & a_- \leq m \\
 (s+1) \quad & a_- \leq m + E + k \cdot E_b + E_b \quad \text{for some fixed } k \geq 0 \\
 (r) \quad & a_- \leq m + E + k \cdot E_b + \ell \cdot E \quad \text{for the same } k \text{ and some fixed } \ell > 0 \\
 (\text{sep}) \quad & 0 \leq -m - E - k \cdot E_b + a_- - 1 \quad \text{for the same } k
 \end{aligned}$$

(0) is true because at s_0 , a is enabled. (s+1) is true because a is enabled one state after s . (r) is true because a is enabled at r ; and $\ell > 0$ because of the $+$. Finally, (sep) is true because q disables a at state s . Adding (s+1)+(sep) gives $1 \leq E_b$. Adding (0)+(sep) gives $1 \leq -E - k \cdot E_b$, and using also $1 \leq E_b$ gives $1 \leq -E - k \cdot E_b \leq -E$. Adding (r)+(sep) gives $1 \leq \ell \cdot E$, contradicting $1 \leq -E$. The system cannot be solved, and no place q separating a at s exists. \square 2

Conjecture 1. A CONVERSE OF PROPOSITION 2

Suppose a word over $\{a, b\}$ is non-PN-solvable and minimal with that property. Then it is (modulo swapping a and b) of the form given in (1). *Strengthened conjecture:* It is either of the form $ab \underbrace{b^j}_{\alpha} b^k ba \underbrace{b^j}_{\alpha} a$ with $j \geq 0, k \geq 1$ or of the

form $aba(ba\alpha)^\ell a$ with $\ell \geq 1$

\square Conj. 1

Proposition 3. ANOTHER SUFFICIENT CONDITION FOR UNSOLVABILITY

Let w be of the form $s_0[\alpha]s[\beta]r[a]$ such that α starts with a and β starts with b .

$$\text{If } \#_a(\beta) \cdot \#_b(\alpha) \geq \#_a(\alpha) \cdot \#_b(\beta) \quad (2)$$

then w is unsolvable.

For instance, for $w = abbaa$, $\alpha = ab$ and $\beta = ba$, and (2) holds true.

Proof: If a place q separates a at s and has marking m at s_0 , then for $E_\alpha = \mathbb{E}(\alpha) = \#_a(\alpha) \cdot E_a + \#_b(\alpha) \cdot E_b$ and $E_\beta = \mathbb{E}(\beta) = \#_a(\beta) \cdot E_a + \#_b(\beta) \cdot E_b$ we have:

$$\begin{aligned} (0) \quad a_- &\leq m && \text{(since } \alpha \text{ starts with } a) \\ (r) \quad a_- &\leq m + E_\alpha + E_\beta \\ (\text{sep}) \quad 0 &\leq -m - E_\alpha + a_- - 1 && \text{(since } \neg s[a]) \end{aligned}$$

Adding (0)+(sep) yields $1 \leq -E_\alpha$, hence (A): $-(\#_a(\alpha)E_a + \#_b(\alpha)E_b) \geq 1$, where E_a and E_b denote the effects of a and b on q , respectively. As before, $E_b \geq 1$. Adding (r)+(sep) yields $1 \leq E_\beta$, hence (B): $(\#_a(\beta)E_a + \#_b(\beta)E_b) \geq 1$. Then,

$$\begin{aligned} -\#_a(\beta) &\geq \#_a(\beta)\#_a(\alpha)E_a + \#_a(\beta)\#_b(\alpha)E_b && \text{(algebra, and by (A))} \\ &\geq \#_a(\beta)\#_a(\alpha)E_a + \#_a(\alpha)\#_b(\beta)E_b && \text{(using (2) and } E_b \geq 1) \\ &\geq \#_a(\alpha) && \text{(algebra, and by (B))} \end{aligned}$$

However, $-\#_a(\beta) \geq \#_a(\alpha)$ implies $\#_a(\beta) = \#_a(\alpha) = 0$, and this is a contradiction since α contains at least one a . Thus, such a place q does not exist. \square 3

Words in which the letters a and b strictly alternate are easy to solve. Therefore, it stands to reason to investigate cases in which a letter occurs twice in a row.

Proposition 4. SOLVABLE WORDS STARTING WITH a CAN BE PREFIXED BY a
If a word av is PN-solvable then aav is, too.

Proof: Let $N = (P, \{a, b\}, F, M_0)$ be a net solving av . We shall construct a net which solves aav . The idea is to obtain such a net by “unfiring” a once from the initial marking of N . Since this may lead to a non-semipositive marking which we would like to avoid, we will first normalise and modify the net N , obtaining another solution N' of av , and then construct a solution N'' for aav (cf. Fig. 4).

For normalisation, we assume that there are two places p_b and q_a ; the first prevents b explicitly in the initial phase, and the second prevents a after the last occurrence of a . They are defined by $M_0(p_b) = 1$, $F(a, p_b) = 1$, $F(b, p_b) = \ell + 1 = F(p_b, b)$, where ℓ is the number of a before the first b in av , and $M_0(q_a) = k$, $F(q_a, a) = 1$, where k is the number of a in av . (All other F values = 0.)

Let $NUF(a) = \{p \in a^\bullet \mid M_0(p) < F(a, p)\}$ be the set of places which do not allow the “unfiring” of a at M_0 . Note that neither p_b nor q_a are in $NUF(a)$. Note also that for every $p \in NUF(a)$, $F(p, a) \leq M_0(p) < F(a, p)$ – the first because a is initially enabled, the second by $p \in NUF(a)$. That is, a has a positive effect on p . Without loss of generality, b has a negative effect on p (otherwise, thanks to the normalising place p_b , p could be deleted without changing the behaviour of N).

For every $p \in NUF(a)$ we add the quantity $F(a, p)$ uniformly to $M_0(p)$, to $F(p, b)$, and to $F(b, p)$, eventually obtaining $N' = (P', \{a, b\}, F', M'_0)$, and we show that N' also solves av . First, both $M_0[a] \wedge \neg M_0[b]$ and $M'_0[a] \wedge \neg M'_0[b]$ (the former by definition, the latter by construction). For an inductive proof, suppose that $M_0[a]M_1[\tau]M$ and $M'_0[a]M'_1[\tau]M'$. We have $M[b]$ iff $M'[b]$ by construction. If

$M[a]$, then also $M'[a]$, since $M \leq M'$. Next, suppose that $\neg M[a]$; then there is some place q such that $M(q) < F(q, a)$. We show that, without loss of generality, $q \notin NUF(a)$, so that q also disables a at M' in N' . If M disables the last a in av , we can take $q = q_a \notin NUF(a)$. If M disables some a which is not the last one in av , then q cannot be in $NUF(a)$, since b acts negatively on such places.

Now, we construct a net $N'' = (P', \{a, b\}, F', M''_0)$ from N' by defining $M''_0(p) = M'_0(p) - F'(a, p) + F'(p, a)$ for every place p . By construction, aav is a firing sequence of N'' . Furthermore, M''_0 does not enable b because of p_b . \square 4

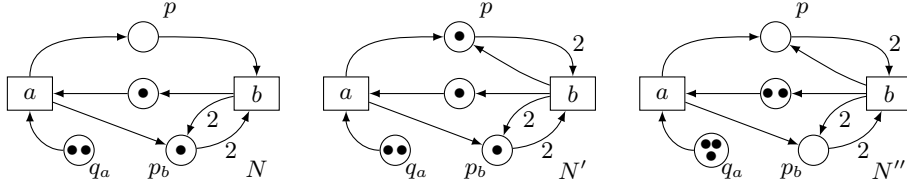


Fig. 4. N is normalised and solves $abab$. N' solves $abab$ as well. N'' solves $aabab$.

As a consequence, if av is minimally PN-unsolvable, then v starts with a b .

Proposition 5. NO aa AND bb INSIDE A MINIMAL UNSOLVABLE WORD

If a minimal non-PN-solvable word is of the form $u = \alpha a$, then either α does not contain the factor aa or α does not contain the factor bb .

Proof: By contraposition. Assume that α contains a factor aa and a factor bb . Two cases are possible:

Case 1: There is a group of a 's which goes after a group of b 's. Let a^m and b^n be such groups, assume that a^m goes after b^n and that there are no groups of a or of b between them. Then u is of the following form

$$s_0 [\dots] q [ab^n (ab)^k a^m] r [\dots]$$

where $n, m \geq 2$, $k \geq 0$. Recombine the letters in u to the following form:

$$s_0 [\dots] q [(ab)b^{n-2}(ba)^{k+1}aa^{m-2}] r [\dots]$$

Since u ends with a , $(ab)b^{n-2}(ba)^{k+1}a$ is a proper subword of u . But it has the form $(abw)b^*(baw)^+a$, with $w = \varepsilon$, which implies its unsolvability by proposition 2, contradicting the minimality of u .

Case 2: All groups of a precede all groups of b . In this case u is of the form

$$aa^{x_0}ba^{x_1} \dots ba^{x_n}b^{y_0}ab^{y_1}ab^{y_2} \dots ab^{y_m}a$$

where at least one of x_i and one of y_j is greater than 1. Consider $\ell = \max\{i \mid x_i > 1\}$. If $\ell = 0$, we get a contradiction to proposition 4. Hence, $\ell > 0$. Let $t = \min\{j \mid y_j > 1\}$. Then u has the form

$$s_0 [a \dots] q [ba^{x_\ell}(ba)^{n-\ell}(ba)^t b^{y_t}] r [\dots a]$$

Recombine the letters in u to the form

$$s_0 [a \dots] q [(ba)a^{x_\ell-2}(ab)^{n-\ell+t+1}bb^{y_\ell-2}] r [\dots a]$$

Hence, u has a proper subword $(ba)a^{x_\ell-2}(ab)^{n-\ell+t+1}b$, which is of the form $(baw)a^*(abw)^+b$ with $w = \varepsilon$, implying its non- PN -solvability, due to proposition 2 with inverted a and b . This again contradicts the minimality of u . \square 5

For these reasons, we are particularly interested in words of the following form:

$$\text{either } ab^{x_1}a \dots ab^{x_n}a \text{ or } b^{x_1}a \dots ab^{x_n} \quad \text{where } x_i \geq 1 \text{ and } n > 1 \quad (3)$$

In the first form, there are no factors aa . If factors bb are excluded and the word starts and ends with an a , then we get words that are of the second form, except for swapping a and b . This swapping is useful in order to understand how words of the two forms are interrelated.

Conjecture 2. A CONVERSE OF PROPOSITION 3

If a word is of the form $w = \alpha\beta a$ where α starts with a and β starts with b , and if w is minimal non- PN -solvable and also of the form given in (3), then inequation (2) holds. *A stronger variant of this conjecture:* If $w = \alpha\beta a$ is of the form

$$w = [\underbrace{ab^{x_1}a \dots ab^{x_{k-1}}}_{\alpha}] s [\underbrace{ba \dots ab^{x_n}}_{\beta} a] \quad \text{with } n \geq 3 \text{ and } x_i \geq 1$$

then a is not separated at state s iff $\#_a(\beta) \cdot \#_b(\alpha) = \#_a(\alpha) \cdot \#_b(\beta)$. \square Conj. 2

5 Some results about words of the form $b^{x_1}a \dots ab^{x_n}$

5.1 Side-places in words of the form $b^{x_1}a \dots ab^{x_n}$

If a word $w = b^{x_1}a \dots ab^{x_n}$ can be solved at all, then side-places may be necessary to do it. However, we will show that in the worst case, only some side-place q around a , preventing a at some state, are necessary.

Lemma 1. SIDE-PLACE-FREENESS AROUND b

If $w = b^{x_1}a \dots ab^{x_n}$ is solvable, then w is solvable without side-place around b .

Proof: Let w and its intermediate states be of the form¹

$$w = s_0[b^{x_1}]s_1[ab^{x_2}]s_2[a] \dots s_{n-1}[ab^{x_n}]s_n \quad (4)$$

Suppose some place p prevents b at some state s_k , for $1 \leq k \leq n-1$. (The only other state at which b must be prevented is state s_n , but that can clearly be done

¹ A note on convention: in the following, we use the letter s to denote states at which b has to be prevented, and p for places doing this. Similarly, we use the letter r to denote states at which a has to be prevented, and q for places doing this.

by a non-side-place, e.g. by an incoming place of transition b that has initially $\sum_{i=1}^n x_i$ tokens.) Note that $b_- > b_+$, because place p allows b to be enabled at the state preceding s_k but not at s_k . Similarly, $a_- < a_+$, because b is not enabled at state s_k but at the immediately following state, which is reached after firing a . From the form (4) of w , we have

$$\begin{aligned}
b_+ &\leq m + x_1(b_+ - b_-) \\
b_+ &\leq m + (x_1 + x_2)(b_+ - b_-) + (a_+ - a_-) \\
&\dots \\
b_+ &\leq m + (x_1 + \dots + x_n)(b_+ - b_-) + (n - 1)(a_+ - a_-) \\
0 &\leq -m - (x_1 + \dots + x_k)(b_+ - b_-) - (k - 1)(a_+ - a_-) + b_- - 1
\end{aligned} \tag{5}$$

The first n inequations assert the semipositivity of the marking of place p (more precisely, its boundedness from below by b_+ , since p may be a side-place) at the n states s_1, \dots, s_n . In our context, if these inequalities are fulfilled, then the marking is $\geq b_+$ at *all* states, as a consequence of $b_- \geq b_+$, $a_- \leq a_+$, and the special form of the word. The last inequality comes from $\neg(s_k[b])$.

We certainly have $0 \leq b_+ \leq b_- \leq m$, because of $b_- \geq b_+$ as noted above, and because b is initially enabled. If $b_+ = 0$, then p is not a side-place around b , and there is nothing more to prove (for p). If $b_+ \geq 1$, we consider the transformation

$$b'_+ = b_+ - 1 \text{ and } b'_- = b_- - 1 \text{ and } m' = m - 1$$

The relation $0 \leq b'_+ \leq b'_- \leq m'$ still holds for the new values. Also, all inequalities in (5) remain true for the new values: in the first n lines, 1 is subtracted on each side, and on the last line, the increase in $-m$ is offset by the decrease in b_- . Thus, we get a solution preventing b with a ‘smaller’ side-place, and we can continue until eventually b_+ becomes zero.

A side-place around b might, however, still be necessary to prevent a at some state. We show next that such side-places are also unnecessary.

Let w and its intermediate states be of the form

$$[b^{x_1-1}\rangle r_1 [b a b^{x_2-1}\rangle r_2 [b a \rangle \dots [a b^{x_k-1}\rangle r_k [b\rangle s_k [a \rangle \dots [b^{x_{n-1}-1}\rangle r_{n-1} [b a b^{x_n}\rangle$$

Suppose some place q as on the right-hand side of figure 2 prevents a at state r_k , for $1 \leq k \leq n - 1$. Symmetrically to the previous case, we have $b_+ > b_-$. This is true because, while q does not have enough tokens to enable a at state r_k , it must have enough tokens to enable a at the directly following state (which we may continue to call s_k). But we also have (w.l.o.g.) $a_+ < a_-$. For $k \geq 2$, this follows from the fact that if the previous a (enabled at the state s_{k-1} just after r_{k-1}) acts positively on q , then q also has sufficiently many tokens to enable a at state r_k . For $k = 1$, it is possible to argue that $a_+ \leq a_-$ is valid without loss of generality. For suppose that q disables a only at r_1 and nowhere else. (This is no loss of generality because for the other states r_k , $k \geq 2$, copies of q can be used.) Then we may consider q' which is an exact copy of q , except that $a_+ = a_- - 1$ for q' . This place q' also disables a at state r_1 (because it has the same marking as

q). Moreover, it does not disable a at any other state after r_1 because it always has $\geq a_- - 1$ tokens, and after the next b , $\geq a_-$ tokens, since $b_+ > b_-$.

Because of $b_+ \geq b_-$ and $a_+ \leq a_-$, q also prevents a at all prior states in the same group of b 's. Moreover, in the last (i.e. n 'th) group of b 's, a can easily be prevented side-place-freely. For place q with initial marking m , we have

$$\begin{aligned}
a_+ &\leq m + x_1(b_+ - b_-) + (a_+ - a_-) \\
a_+ &\leq m + (x_1 + x_2)(b_+ - b_-) + 2(a_+ - a_-) \\
&\dots \\
a_+ &\leq m + (x_1 + \dots + x_{n-1})(b_+ - b_-) + (n-1)(a_+ - a_-) \\
0 &\leq -m - (x_1 + \dots + x_k - 1)(b_+ - b_-) - (k-1)(a_+ - a_-) + a_- - 1
\end{aligned} \tag{6}$$

The first $n-1$ inequations assert the semipositivity of the marking of place q (more precisely, its boundedness from below by a_+ , since q may be a side-place of a) at the $n-1$ states r_1, \dots, r_{n-1} . If these inequalities are fulfilled, then the marking is $\geq a_+$ at *all* states after the first a , as a consequence of $b_+ \geq b_-$ and the special form of the word. The last inequality asserts that place q prevents transition a at state r_k , hence effects the event/state separation of a at r_k .

If b_- is already zero, place q is not a side-place of b . Otherwise, we may perform the transformation

$$b'_+ = b_+ - 1 \text{ and } b'_- = b_- - 1 \text{ and } m' = m$$

because of $b_+ \geq b_-$ as noted above. The left-hand sides of the first $n-1$ inequalities in (6) do not decrease, and neither do the right-hand sides. The same is true for the last inequality. \square 1

Lemma 2. SIDE-PLACE-FREENESS AROUND a , PREVENTING b

Suppose $w = b^{x_1} a b^{x_2} a \dots a b^{x_n}$. If w is solvable by a net in which some place p separates b , then we may w.l.o.g. assume that p is not a side-place around a .

Proof: The equation system (5) is invariant under the transformation

$$a'_+ = a_+ - 1 \text{ and } a'_- = a_- - 1 \text{ and } m' = m$$

as neither the left-hand sides nor the right-hand sides change their values. \square 2

If some place q prevents transition a , then a side-place q connected to a may be present. It may not always be possible to remove such a side-place. Consider, for instance, the word $w = bbbabab$. It is of the form (4), and any net solving $bbbabab$ necessarily contains a side-place around transition a .² The word $bbabbababab$ can also not be solved without a side-place (but $bbabbabab$ can). So far, no tight (weak) sufficient conditions for solvability, or solvability without side-places around a , are known. However, the next lemma shows that the presence of a side-place around a may be due to there being ‘‘many’’ initial b 's. That is, if x_1 is ‘‘small enough’’, then such a side-place is not necessary.

² The reader is invited to use [7] in order to verify this claim; we will not include a proof in this paper for lack of space.

Lemma 3. SIDE-PLACE-FREENESS AROUND a , PREVENTING a

Suppose $w = b^{x_1} a b^{x_2} a \dots a b^{x_n}$. If $x_1 \leq \min\{x_2, \dots, x_{n-1}\}$ and if w is solvable by a net in which some place q prevents transition a at state r_k with $1 \leq k \leq n$, then we may w.l.o.g. assume that q is not a side-place around a .

Proof: For preventing a at state r_n , we only need a place with no input and output a (weight 1) which has $n - 1$ tokens initially.

Suppose q prevents a at state r_k , with $1 \leq k \leq n - 1$. From previous considerations, we know $a_+ < a_-$ and $b_+ > b_-$, and we may assume, w.l.o.g., that q is not a side-place around b , i.e., that $b_- = 0$. The initial marking m of q and the remaining arc weights a_+, a_-, b_+ satisfy the system of inequations (6), except that it is simplified by $b_- = 0$. If $a_+ = 0$, then q is already of the required form. For $a_+ > 0$, we distinguish two cases.

Case 1: $m > 0$ and $a_+ > 0$. Then consider the transformation

$$m' = m - 1 \text{ and } a'_+ = a_+ - 1 \text{ and } a'_- = a_- - 1$$

By $m > 0$ and $a_- \geq a_+ > 0$, we get new values $m', a'_+, a'_- \geq 0$. Moreover, (6) remains invariant under this transformation. So, q' serves the same purpose as q , and it has one incoming arc from a less than q . By repeating this procedure, we either get a place which serves the same purpose as q , or we hit Case 2.

Case 2: $m = 0$ and $a_+ > 0$. In this case, we consider the transformation

$$m' = m = 0 \text{ and } a'_+ = 0 \text{ and } a'_- = a_-$$

Such a transformation also guarantees $m', a'_+, a'_- \geq 0$. Also, the last line of (6) is clearly satisfied with these new values, since the value of its right-hand stays the same (for $k = 1$) or increases (for $k > 1$). To see that the first $n - 1$ lines of (6) are also true with the new values (and with $b_- = 0$), and that we can, therefore, replace q by q' , we may argue as follows. At any marking \tilde{m} reached along the execution of w , we have the following:

$$\tilde{m}(q) \geq \tilde{m}(q') \geq 0 \tag{7}$$

These inequalities imply that the new place q' prevents a at r_k , whenever the old one, q , does, and that, moreover, no occurrences of a are excluded by the place q' where they should not be prohibited.

The first of the inequalities (7) holds because it holds initially (when $\tilde{m} = m$, then $\tilde{m}(q) = m = m' = \tilde{m}(q')$), and because the effect of a before the transformation is $(a_+ - a_-)$, and after the transformation, it is $(-a_-)$. In other words, a reduces the token count on q' more than it does so on q , while b has the same effect on q' as on q . To see the second inequality in (7), let $x = \min\{x_2, \dots, x_{n-1}\}$. Then

$$a_- \leq x_1 \cdot b_+ \leq x \cdot b_+$$

The first inequality follows because $m = 0$ and q has enough tokens after the first x_1 occurrences of b in order to enable a . The second inequality follows from

$x_1 \leq x$. But then, since a only removes a_- tokens from q' and the subsequent block of b 's puts at least $x \cdot b_+$ tokens back on q' , the marking on q' is always ≥ 0 , up to and including the last block of b 's. \square 3

Corollary 1. SIDE-PLACE-FREE SOLVABILITY WITH FEW INITIAL b 'S

If $w = b^{x_1}ab^{x_2}a \dots ab^{x_n}$ is solvable, then side-places are necessary, at worst, between a and q , where q is some place preventing a at one of the states r_k with $1 \leq k < n - 1$. If $w = b^{x_1}ab^{x_2}a \dots ab^{x_n}$ is solvable and $x_1 \leq \min\{x_2, \dots, x_{n-1}\}$, then w is solvable side-place-freely. \square 1

5.2 Solving words aw from words of the form $w = b^{x_1}a \dots ab^{x_n}$

Solving a word of the form $w = b^{x_1}a \dots ab^{x_n}$ side-place-freely allows us to draw some conclusion about prepending a letter a to it, as follows.

Lemma 4. PREVENTING a IN aw

Suppose $w = b^{x_1}ab^{x_2}a \dots ab^{x_n}$ is solvable side-place-freely. Then in aw , all occurrences of a can be separated side-place-freely.

Proof: In order to prevent a in w side-place-freely at any state r_k , the system (6) has a solution with $a_+ = 0$ and $b_- = 0$ for any fixed $1 \leq k \leq n - 1$. This refers to a pure input place q of a , which may or may not be an output place of b . In order to prevent a in aw side-place-freely, we need to consider the states r_k as before (but shifted to the right by one index position, still just before the last b of the k 'th group of b 's) and a correspondingly modified system as follows:

$$\begin{aligned} 0 &\leq m' + (x_1 + \dots + x_i) \cdot (b'_+) + (i + 1) \cdot (-a'_-) & \text{for all } 0 \leq i \leq n - 1 \\ 0 &\leq -m' - (x_1 + \dots + x_k - 1) \cdot (b'_+) - k \cdot (-a'_-) + a'_- - 1 \end{aligned} \quad (8)$$

where m' , b'_+ and a'_- refer to a new pure place q' preventing a at state r_k in aw . The line with $i = 0$ was added because m' is required to be bounded from below and a must be enabled initially. (In (6), nonnegativity of m follows from the line with $i = 1$ and b being the first transition of w , which is no longer true in aw .) Consider the transformation

$$m' = m + a_- \text{ and } b'_+ = b_+ \text{ and } a'_- = a_-$$

These values satisfy (8), provided m , b_+ and a_- (together with $a_+ = 0$ and $b_- = 0$) satisfy (6). The line with $i = 0$ follows from $m' = m + a_- \geq 0$. The other lines corresponding to $i \geq 1$ reduce to the corresponding lines in (6), since the additional $(-a_-)$ at the end of each line is offset by the additional $(+a_-)$ at the beginning of the line. The last line (which belongs to state r_k at which a is separated) corresponds to the last line of (6), because the decrease by a_- at the beginning of the line is offset by an increase by a_- in the term $k \cdot (-a'_-)$ (compared with $(k - 1) \cdot (-a_-)$ as in (6)). \square 4

Note 1: In order to disable a at r_k , q could be replaced by a place q' obtained by duplicating q and changing the initial marking m to $m' = m + a_-$. Intuitively, this means that m' is computed from m by “unfiring” a once.

Note 2: Place q should not be removed as soon as q' is added, because q could also be preventing a at some other r_k . In that case, a new place q'' must be computed from q for this different value of k . We may forget about q only after all the relevant indices k have been processed.

Next, consider an input place p of b in a side-place-free solution of w and suppose that p prevents b at state s_k . Suppose that we want to solve aw . If p is not also an output place of a , then it can simply be retained unchanged, and with the same marking, prevent b at corresponding states in aw and in w . However, if p is also an output place of a , there may be a problem, because “unfiring” a in the initial marking may lead to negative tokens on p . This is illustrated by the example $babbabb$ which has a side-place-free solution, as shown on the left-hand side of figure 5.

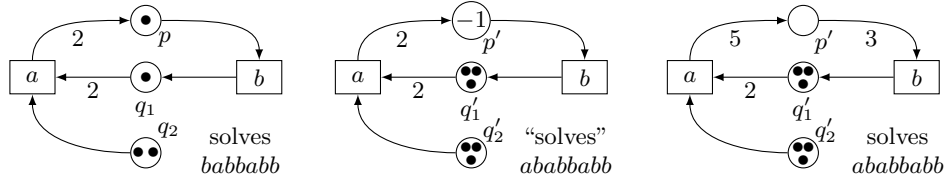


Fig. 5. Solving $babbabb$ (l.h.s.), (almost) $ababbabb$ (middle), and $ababbabb$ (r.h.s.).

The places q_1, q_2 can be treated as in the above proof (that is, by changing their markings by “unfiring” a) and yield new places q'_1, q'_2 with marking $\{(q'_1, 3), (q'_2, 3)\}$. If we allowed negative markings, then a new place p' with initial marking $(p', -1)$ (and otherwise duplicating p) would do the job of solving $ababbabb$ (as in the middle of the figure). However, we shall need a more refined argument in order to avoid negative markings.

Let p' be a general new place which is supposed to prevent b at state s_k in aw . In order to check the general solvability of aw if w is side-place-freely solvable, we consider a general transformation

$$m' = m + \mu \quad , \quad b'_+ = b_+ + \beta_+ \quad , \quad b'_- = b_- + \beta_- \quad , \quad a'_+ = a_+ + \alpha_+ \quad , \quad a'_- = a_- + \alpha_-$$

where $\mu \geq -m$, $\beta_+ \geq -b_+$, $\beta_- \geq -b_-$, $\alpha_+ \geq -a_+$ and $\alpha_- \geq -a_-$, as well as a new inequation system:

$$\begin{aligned} b'_+ &\leq m' + (x_1 + \dots + x_i) \cdot (b'_+ - b'_-) + i \cdot (a'_+ - a'_-) \quad \text{for } 1 \leq i \leq n \\ 0 &\leq -m' - (x_1 + \dots + x_k) \cdot (b'_+ - b'_-) - k \cdot (a'_+ - a'_-) + b'_- - 1 \end{aligned}$$

This system has to be compared with a restricted form of (5) (setting $b_+ = a_- = 0$, since the solution of w is pure). Doing this by line-wise comparison, we get the

following inequation system for the new value differences:

$$\begin{aligned}
\mu &\geq -m, \beta_+ \geq -b_+, \beta_- \geq -b_-, \alpha_+ \geq -a_+, \alpha_- \geq -a_- \\
\beta_+ &\leq \mu + (x_1 + \dots + x_i) \cdot (\beta_+ - \beta_-) + i \cdot (\alpha_+ - \alpha_-) + a_+ \\
0 &\leq -\mu - (x_1 + \dots + x_k) \cdot (\beta_+ - \beta_-) - k \cdot (\alpha_+ - \alpha_-) - a_+ + \beta_-
\end{aligned} \tag{9}$$

The lines with i must be solved simultaneously for every $1 \leq i \leq n$ while the line with k must be solved individually for every $1 \leq k \leq n - 1$, in order to get a place preventing b at state s_k . This leads to the following lemma.

Lemma 5. SOLVING aw FROM w

Suppose $w = b^{x_1}ab^{x_2}a \dots ab^{x_n}$ is solvable side-place-freely. Then aw is solvable.

Proof: Suppose that a pure place p with parameters b_- (arc into b), a_+ (arc from a) and m (initial marking) is given and suppose it separates b from s_k in w . This place solves (5) for that particular k . We distinguish two cases:

Case 1: $a_+ \leq m$. In this case, the place p can essentially be re-used for the same purpose in the solution (that we construct in this way) for aw , since (9) is solved by putting

$$\mu = -a_+, \beta_+ = \beta_- = 0, \alpha_+ = \alpha_- = 0$$

Hence, a place p' which differs from p only by its initial marking ($m' = m - a_+$ instead of m) separates b at s_k in aw .

Case 2: $a_+ > m$. In this case, (9) can be solved by

$$\mu = -m, \beta_+ = \beta_- = a_+ - m, \alpha_+ = \alpha_- = 0$$

That is, we may replace p by a place p' with zero initial marking and adding uniformly the value $a_+ - m$ to the incoming and outgoing arcs of b , creating a side-place around b . \square 5

For instance, in the solution of $babbabb$ shown on the left-hand side of figure 5, the place p from a to b satisfies $m=1$, $b_-=1$, $b_+=0$, $a_-=0$ and $a_+=2$. (9) is solved by $\mu = -1$, $\beta_- = 2$, $\beta_+ = 0$, $\alpha_- = 0$ and $\alpha_+ = 3$. Hence with $m'=m-1$, $b'_-=b_-+2$, $b'_+=b_+$, $a'_-=a_-$ and $a'_+=a_++3$, the net shown on the right-hand side of figure 5 is a pure solution of $ababbabb$. (Place p' prevents b not only in states s_1 and s_2 but also in the initial state and in the final state.) There exist words such as $w_1 = bbbabab$ or $w_2 = bbabbababab$, however, which can be solved but for which aw is not solvable. We have a converse of Lemma 5:

Lemma 6. SOLVING w SIDE-PLACE-FREELY FROM aw

If aw has a solution, then w has a side-place-free solution.

Proof: Suppose that aw has a solution in which some place q' , preventing a , has a side-place around a . Because q' prevents a , $a'_- > a'_+$ (unless it is the first a , but then we don't need q' in solving w). Because a is enabled initially, $m' \geq a'_-$. But then, the transformation $a''_- = a'_- - a'_+$, $a''_+ = 0$, $m'' = m' - a'_+$ yields another place q'' which is not a side-place around a but serves the same purpose as q' . The rest of the proof follows because the above transformations (removing side-places around b , or side-places around a which prevent b) do not introduce any new side-places around a . \square 6

Corollary 2. SIDE-PLACE-FREE SOLVABILITY OF $b^{x_1}ab^{x_2}a \dots ab^{x_n}$
 $w = b^{x_1}ab^{x_2}a \dots ab^{x_n}$ is solvable side-place-freely iff aw is solvable. \square 2

6 Concluding remarks

If the characterisations of minimal Petri net solvable binary words proposed in this paper are valid, then the usual linear solver for detecting them can be replaced by a pattern-matching algorithm based on conjecture 1, or by a letter-counting algorithm based on conjecture 2. We have described a variety of results providing some insight into this class of words. There are several other facts about them which we did not have space to describe. For example, if a word is solvable side-place-freely, then so is the reverse word. Also, if a word is solvable, then it is solvable by places having exactly one outgoing transition. (This property is not shared by words with three or more letters, a counterexample being $abcbaa$.) Moreover, PN-solvable words are balanced in the following sense. Referring to $w = b^{x_1}ab^{x_2}a \dots ab^{x_n}$, call w balanced if there is some x such that $x_i \in \{x, x+1\}$ for all $2 \leq i \leq n-1$. We can prove that if $w = b^{x_1}ab^{x_2}a \dots ab^{x_n}$ is PN-solvable, then w is balanced, and moreover, $x_n \leq x+1$. Presenting these, and other, properties of PN-solvability must however be left to future publications.

Acknowledgements: We would like to thank Harro Wimmel and Uli Schlachter for valuable comments, and Uli for finding weird words (amongst them, $abcbaa$ as mentioned in the conclusion) using [7, 3]. We thank also the anonymous reviewers for their remarks which allowed to improve the presentation of the paper.

References

1. É. Badouel, P. Darondeau: Theory of Regions. In W. Reisig, G. Rozenberg (eds): Lectures on Petri Nets I: Basic Models. LNCS Vol. 1491, 529–586 (1998).
2. E. Best, R. Devillers: State Space Axioms for T-Systems. G. Lüttgen, F. Corradini (eds): Special volume on the occasion of Walter Vogler's 60th birthday. Acta Informatica, Vol. 52(2-3), 133-152 (2015).
3. E. Best, U. Schlachter: Analysis of Petri Nets and Transition Systems. ICE'2015.
4. B. Caillaud: <http://www.irisa.fr/s4/tools/synet/>
5. T. Murata: Petri Nets: Properties, Analysis and Applications. Proc. of the IEEE, Vol. 77(4), 541-580 (1989).
6. M. Piątkowski et al.: <http://folco.mat.umk.pl/unsolvable-words> (2015).
7. U. Schlachter et al.: <https://github.com/Cv0-Theory/apt> (2013).