Bouziane's transformation of the Petri net reachability problem and incorrectness of the related algorithm

Petr Jančar

Dept. of Computer Science, Technical University of Ostrava (FEI-VŠB),
17. listopadu 15, CZ - 708 33 Ostrava, Czech Republic
(e-mail: Petr.Jancar@vsb.cz)

Abstract. Proceedings of FOCS'98 contain a paper by Zakariae Bouziane, who sketches a new representation of the Petri net reachability problem and claims to provide a new algorithm solving the problem. In this note, the idea of Bouziane's approach is explained, and a serious flaw of the algorithm is exposed.

1 Introduction

The Petri net reachability problem (PNRP), which asks if a given state (i.e., marking) is reachable from the initial state in a given (place/transition) Petri net, is a well-known problem also outside the Petri net community. It is surprisingly tricky from the complexity point of view. Its decidability had been a challenging open problem for about a decade, starting in the early 1970s. Positive solutions for some subcases were given by several researchers, and then a "solution" for the general case was presented at STOC'77 – but it turned out to be incomplete afterwards. The decidability was finally established by E.W. Mayr and presented at STOC'81 [4, 5]. He used an involved combination of nice ideas and developed an algorithm with nonprimitive recursive complexity. But since it was a nontrivial task to verify Mayr's involved proof, some other researchers have further elaborated on it, aiming at improving its general understandability; the first one was Kosaraju at STOC'82. Nevertheless nobody has improved the known upper bound on the complexity, which thus remains nonprimitive recursive. On the other hand, PNRP is known to be EXPSPACE-hard due to Lipton [3] (the proof can be also found, e.g., in [2]). Thus the exact complexity of PNRP remains unclear and keeps to constitute an intellectual challenge.

In the Proceedings FOCS'98 [1], Z. Bouziane claims to provide a new algorithm for PNRP; at first sight it looks conceptually simpler than Mayr's algorithm, and its complexity is claimed to be primitive recursive (double-exponential, in fact). Bouziane first transforms PNRP into an equivalent problem in a different setting. This new problem, let us denote it by BP, might resemble a special (and decidable) version of the Post correspondence problem, very roughly sketched as follows: for given pairs $(u_0, v_0), (u_1, v_1), (u_2, v_2), \ldots (u_n, v_n)$, where each u_i (v_i) represents a tuple of sets of numbers, it asks if there is a finite sequence $i_1, i_2, \ldots, i_m$ of indices such that $u_0 u_{i_1} u_{i_2} \ldots u_{i_m}$ and $v_0 v_{i_1} v_{i_2} \ldots v_{i_m}$ represent (tuples of) sets of numbers which are linearly related in a certain sense.

Since BP is an alternative of PNRP, formulated in a different framework, it surely can be worth to try to tackle PRNP by studying BP. Bouziane did this and sketched an algorithm for BP, based on finite automata constructions. Unfortunately, the paper in FOCS'98 Proceedings was written in a very unclear form and it raised serious questions concerning its validity. Nevertheless, the author (i.e., Z. Bouziane) has never produced any more elaborated version, and he seems to have abandoned this research.

The aim of this note is to clarify the transformation from PNRP to BP – which is valid but hardly understandable from [1] – and then to expose a serious flaw of Bouziane's algorithm. The note can thus also serve for attracting a new attention to an interesting and challenging problem.

In Section 2 we define PNRP and give our running example. In Section 3 we explain the transformation of PNRP to the equivalent BP. Section 4 then exposes the flaw of Bouziane's algorithm.

2 Petri net reachability problem

We now state PNRP and give a simple example on which Bouziane's algorithm will be later contradicted.

Definition 1 *A* Petri net *is a tuple $N = (P, T, F)$ where $P = \{p_1, p_2, \dots, p_m\}$ and $T = \{t_1, t_2, \dots, t_n\}$ are finite disjoint sets of* places *and* transitions *respectively and $F : (P \times T) \cup (T \times P) \to \{0, 1\}$ is a flow function, i.e., the characteristic function of a set of* arcs*.*

A marking *M of the net N is a mapping $M : P \to \mathbb{N}$ associating a nonnegative number of* tokens *to each place; we implicitly assume an ordering of places and view M as a vector from $\mathbb{N}^m$. Addition, subtraction and ordering $\leq$ on $\mathbb{N}$ are extended to vectors component-wise.*

For a transition t, we define ${}^\bullet t = \big(F(p_1, t), F(p_2, t), \dots, F(p_m, t)\big)$ and $t^\bullet = \big(F(t, p_1), F(t, p_2), \dots, F(t, p_m)\big)$.

We put $M \xrightarrow{t} M'$ (marking M changes to M' by performing transition t) iff $M \geq {}^\bullet t$ and $M' = M - {}^\bullet t + t^\bullet$; in the natural way, we extend the definition to $M \xrightarrow{w} M'$ where w is a finite sequence of transitions.

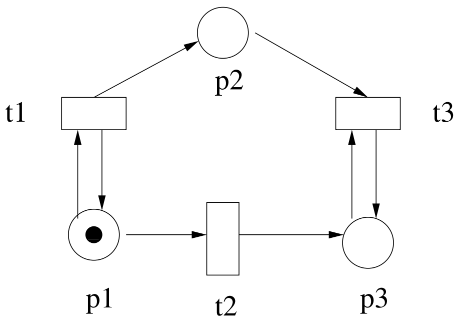

Figure 1:

Figure 1 shows an example of a net with 3 places (circles) and 3 transitions (rectangles); also the marking $(1, 0, 0)$ is depicted. We can check, e.g., that ${}^\bullet t_1 = (1, 0, 0)$, $t_1^\bullet = (1, 1, 0)$, and that $(1, 0, 0) \xrightarrow{w} (0, 2, 1)$ for $w = t_1 t_1 t_1 t_2 t_3$.

Definition 2 *Petri net reachability problem (PNRP):*
Instance: *a Petri net N and two its markings M_0, M_f*
Question: *is there a sequence w of transitions, called a* witness sequence*, such that $M_0 \xrightarrow{w} M_f$?*

If we put $M_0 = (1, 0, 0)$ and $M_f = (0, 0, 1)$ in our example net on Fig. 1, we can observe that the witness sequences are precisely the sequences $w = (t_1)^q t_2 (t_3)^q$, where $q \in \mathbb{N}$ (t^q means $tt \dots t$ where t appears q times).

It is also useful to observe that we can generally confine ourselves to the "no-cycle" witness sequences – these pass through any marking at most once. (If $M_0 \xrightarrow{u} M \xrightarrow{v} M \xrightarrow{z} M_f$ then also $M_0 \xrightarrow{uz} M_f$.)

3 Transformation of PNRP to BP

Bouziane transforms the problem of finding if there is a witness sequence to an equivalent problem of finding if there is a collection of finite subsets of $\mathbb{N}$ satisfying certain conditions. In fact, it is described unclearly in [1], so we try to explain this more precisely here. It serves to one "negative" aim, namely to demonstrate the incorrectness of Bouziane's algorithm, but also to a positive aim – to enable future possible using of this transformation in clarifying the complexity of PNRP.

For informal explanations, we use the example net from Fig. 1 where we put $M_0 = (1,0,0)$ and $M_f = (0,0,1)$.

Definition 3 *Given $m \in \mathbb{N}$, ϕ denotes the "Gödel coding" $\phi : \mathbb{N}^m \to \mathbb{N}$ defined as $\phi(x_1, x_2, \ldots, x_m) = prime_1^{x_1} prime_2^{x_2} \ldots prime_m^{x_m}$ where $prime_i$ is the i-th prime number.*

Given a Petri net, with an ordering on its m places, the meaning of $\phi(M)$ for a marking M and of $\phi({}^\bullet t)$, $\phi(t^\bullet)$ for a transition t is thus induced.

Note that $M \xrightarrow{t} M'$ iff

- $\phi(M)$ is divisible by $\phi({}^\bullet t)$, and
- $\phi(M') = \phi(M) \cdot \delta(t)$, where $\delta(t) = \phi(t^\bullet)/\phi({}^\bullet t)$ (number $\delta(t)$ is rational).

Note also that the conditions imply that $\phi(M')$ is divisible by $\phi(t^\bullet)$.

For brevity we further say just "a marking" instead of "the (Gödel) code of an marking" etc. So in the example, the (code of the) initial marking is 2, the final marking is 5. Consider one particular witness sequence and the corresponding path through markings:

$$2 \xrightarrow{t_1} 6 \xrightarrow{t_1} 18 \xrightarrow{t_1} 54 \xrightarrow{t_2} 135 \xrightarrow{t_3} 45 \xrightarrow{t_3} 15 \xrightarrow{t_3} 5$$

Let us now distribute the markings of this path into classes $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n$, so that in $\mathcal{P}_i$ ($i = 1, 2, \ldots, n$) are precisely those markings which are entered by transition t_i. Hence in our case $n = 3$ and $\mathcal{P}_1 = \{6, 18, 54\}, \mathcal{P}_2 = \{135\}, \mathcal{P}_3 = \{5, 15, 45\}$. (See Fig. 2, ignoring the horizontal "cut" lines at the moment.) Similarly we distribute the markings to classes $\mathcal{P}_{n+1}, \mathcal{P}_{n+2}, \ldots, \mathcal{P}_{2n}$, so that in $\mathcal{P}_{n+i}$ ($i = 1, 2, \ldots, n$) are precisely those markings which are left by transition t_i. Hence in our case $\mathcal{P}_4 = \{2, 6, 18\}$, $\mathcal{P}_5 = \{54\}$, $\mathcal{P}_6 = \{15, 45, 135\}$.

We have thus given an example of a witness collection defined by the following two definitions. Here $A \cdot b$, where $A \subseteq \mathbb{N}$ and b is a (rational) number, we mean the set $\{\, x \cdot b \mid x \in A \,\}$.

Definition 4 *Given a $(2n+2)$-tuple of natural numbers $(\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n, init, fin)$, a* collection $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_{2n}$ *of subsets of $\mathbb{N}$ is* successful *(wrt the tuple) iff the following two conditions hold:*

(1) *a/* *$\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n$ are pairwise disjoint and doo not contain $init$, similarly $\mathcal{P}_{n+1}, \mathcal{P}_{n+2}, \ldots, \mathcal{P}_{2n}$ are pairwise disjoint and do not contain fin*

b/ $\{init\} \cup \mathcal{P}_1 \cup \mathcal{P}_2 \cup \ldots \cup \mathcal{P}_n = \{fin\} \cup \mathcal{P}_{n+1} \cup \mathcal{P}_{n+2} \cup \ldots \cup \mathcal{P}_{2n}$

c/ *each element of $\mathcal{P}_i$ ($i = 1, 2, \ldots, n$) is divisible by α_i, each element of $\mathcal{P}_{n+i}$ ($i = 1, 2, \ldots, n$) is divisible by β_i*

(2) $\mathcal{P}_i = \mathcal{P}_{n+i} \cdot (\alpha_i/\beta_i)$ *for each* $i = 1, 2, \ldots, n$

Definition 5 *Given a Petri net, with transitions $t_1, t_2, \ldots, t_n$, and its markings M_0, M_f, by a* witness collection *we mean any collection $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_{2n}$ of finite subsets of $\mathbb{N}$ which is successfull wrt $\alpha_1 = \phi(t_1^\bullet), \alpha_2 = \phi(t_2^\bullet), \ldots, \alpha_n = \phi(t_n^\bullet)$, $\beta_1 = \phi({}^\bullet t_1), \beta_2 = \phi({}^\bullet t_2), \ldots, \beta_n = \phi({}^\bullet t_n)$, $init = \phi(M_0)$, $fin = \phi(M_f)$.*

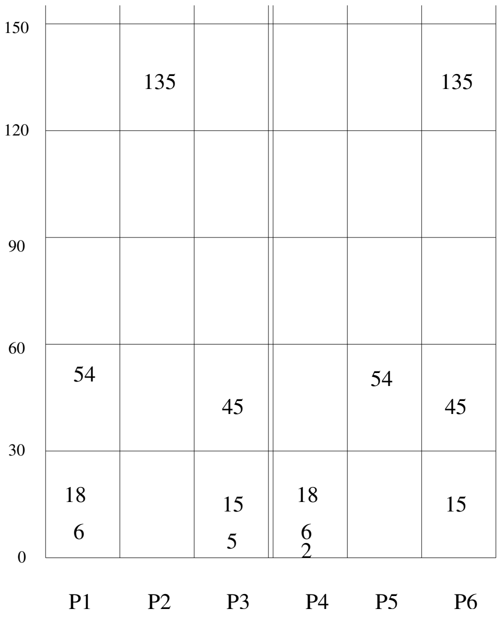

Figure 2:

Generalizing our concrete example, we can easily verify that, given any Petri net and its markings M_0, M_f, any witness sequence translates into a witness collection. On the other hand, we can easily observe that any witness collection "contains" a witness sequence, which can be got by the following "program":

$c_1 := \phi(M_0);\ i := 1$
while $c_i \neq \phi(M_f)$ **do** (find $\mathcal{P}_{n+j}$ s.t. $c_i \in \mathcal{P}_{n+j}$; $c_{i+1} := c_i \cdot \phi(t_j^\bullet)/\phi({}^\bullet t_j);\ i := i+1$)

Note that a witness collection can also contain "separated cycles" (consider adding 10 to $\mathcal{P}_3$ and $\mathcal{P}_4$ and 30 to $\mathcal{P}_1$ and $\mathcal{P}_6$); these "cycles" can even not correspond to (the codes of) markings (consider adding 70 to $\mathcal{P}_3$ and $\mathcal{P}_4$ and 210 to $\mathcal{P}_1$ and $\mathcal{P}_6$). Nevertheless, we have shown

Lemma 6 *For a given Petri net, M_f is reachable from M_0 iff there is a witness collection.*

Remark. An idea of replacing the problem of finding a witness sequence by the problem of finding a collection of finite sets of markings was (indirectly) present, e.g., in [6] (section 3); in fact, Mayr and Meyer deal with polynomials (with m variables) which can naturally represent sets of markings (for m places). Bouziane also uses polynomials, namely one-variable polynomials since he codes markings by integers; but this serves just as a notation and the approach can be presented without using polynomials – as we do in this note.

Let BP denote the problem of deciding whether there is a successful collection wrt a given $(2n+2)$-tuple; we assume the numbers given in binary. We have thus proved:

Theorem 7 *PNRP is polynomially reducible to BP.*

BP is also easily seen to be reducible to PNRP, using prime decomposition of the given numbers. So polynomiality of such a reduction would be questionable but that is not our concern here.

4 Exposing the flaw

After establishing (an equivalent of) Lemma 6, Bouziane shows that the collections which satisfy (1) and need not satisfy (2) have a very regular structure. To demonstrate this, we use some technical definitions, including condition (1^{hom}) which is the "homogeneous" version of (1) (where $init$ and fin are omitted).

Definition 8 *Given a* $(2n{+}2)$*-tuple* $(\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n, init, fin)$*, a* collection $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_{2n}$ *of subsets of* $\mathbb{N}$ *is called a* (1)-collection *iff it satisfies condition (1) of Definition 4.*

It is called a (1^{hom})-collection *iff the following condition (*1^{hom}*) holds:*

a/ $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n$ *are pairwise disjoint, and* $\mathcal{P}_{n+1}, \mathcal{P}_{n+2}, \ldots, \mathcal{P}_{2n}$ *are pairwise disjoint*

b/ $\mathcal{P}_1 \cup \mathcal{P}_2 \cup \ldots \cup \mathcal{P}_n = \mathcal{P}_{n+1} \cup \mathcal{P}_{n+2} \cup \ldots \cup \mathcal{P}_{2n}$

c/ each element of $\mathcal{P}_i$ *(*$i = 1, 2, \ldots, n$*) is divisible by* α_i*, each element of* $\mathcal{P}_{n+i}$ *(*$i = 1, 2, \ldots, n$*) is divisible by* β_i

Definition 9 *Given a* $(2n{+}2)$*-tuple* $(\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_n, init, fin)$*, a number* $\ell_0 \in \mathbb{N}$ *is a* cut number *iff* ℓ_0 *is divisible by all* α_i *and by all* β_i *(*$i = 1, 2 \ldots, n$*) and is greater than both* $init$ *and* fin*. By a* basic segment *(wrt a cut number* ℓ_0*) we mean any (1)-collection whose union is a subset of* $\{0, 1, \ldots, \ell_0{-}1\}$*. By a* basic hom-segment *we mean any (*1^{hom}*)-collection whose union is a subset of* $\{0, 1, \ldots, \ell_0{-}1\}$*. By an* r-th segment *(*$r = 1, 2, \ldots$*) we mean any (*1^{hom}*)-collection whose union is a subset of* $\{r\ell_0, r\ell_0{+}1, \ldots, r\ell_0{+}\ell_0{-}1\}$.

Definition 10 *The* mod-ℓ_0-image *of a set* $A \subseteq \mathbb{N}$ *is the set* $\{\, x \bmod \ell_0 \mid x \in A \,\}$*. The* mod-$\ell_0$-image *of a collection of sets arises by replacing each set in the collection by its mod-*ℓ_0*-image.*

Let us again consider our example, and let us take $\ell_0 = 2 \cdot 3 \cdot 5 = 30$ as a cut number. Any collection $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_{2n}$ can be thought as cut into segments of height ℓ_0; we illustrate it on Fig. 2 for our above given collection. Note that the mod-ℓ_0-images of the 1st, 2nd, 3rd and 4th segment are basic hom-segments – where the mod-ℓ_0-images of the 2nd and 3rd segment are the same (the collection of empty sets in this case). Generally we observe:

Fact 11 *For a cut number* ℓ_0*, any collection whose union is a subset of* $\{r\ell_0, r\ell_0{+}1, \ldots, r\ell_0{+}\ell_0{-}1\}$*,* $r \geq 1$*, is an* r*-th segment iff its mod-*ℓ_0*-image is a basic hom-segment.*

For a $(2n{+}2)$-tuple of numbers and a cut number, we can easily construct the set of all (finitely many) basic segments – and index the set, e.g., by numbers. Similarly we can construct (and index) the set of all (finitely many) basic-hom segments. Using Fact 11, we can easily observe that (1)-collections correspond precisely to the following (1)-words:

Definition 12 *Given a* $(2n{+}2)$*-tuple of numbers and a cut number, a* (1)-word *is a sequence* bu *where* b *is (the index of) a basic segment and* u *is a finite sequence of (the indices of) basic hom-segments.*

Fact 13 *(Given a* $(2n{+}2)$*-tuple of numbers and a cut number,) there is a (straightforward) bijection between the set of (1)-collections and the set of (1)-words.*

Bouziane tries to show that a finite automaton can recognize precisely such (1)-words which correspond to the successful collections (i.e., those satisfying both (1) and (2)). In principle, he tries to prove the following claim; but we will lead this to a contradiction.

Claim 14 (false !) *Given a Petri net and M_0, M_f, we can (effectively) construct a finite automaton which accepts exactly the witness words, i.e., (1)-words corresponding to witness collections.*

If the claim were true, the reachability question would be equivalent to the question if the automaton accepts at least one word, and the size of the automaton could serve for deriving some (elementary) complexity upper bound for the reachability problem; this is what Bouziane did.

But we show that the language of witness words can be nonregular; so there exists no such finite automaton then. It is sufficient to consider our simple example. First note that, in the witness collection corresponding to a witness sequence $(t_1)^q t_2 (t_3)^q$, the sets $\mathcal{P}_2$ and $\mathcal{P}_5$ are singletons, let us denote $\mathcal{P}_2 = \{m_2\}$ and $\mathcal{P}_5 = \{m_5\}$, and they contain the two greatest markings ($m_2 = 5 \cdot 3^q$ and $m_5 = 2 \cdot 3^q$); moreover, $m_2 = \frac{5}{2} m_5$ – as required by (2). Hence the difference between the two greatest markings increases with increasing q. Now assume that a relevant automaton A exists, and denote the number of its states by r. We can take a witness sequence $(t_1)^q t_2 (t_3)^q$ for sufficiently large q so that the difference between the two greatest markings is greater than $\ell_0 \cdot (r{+}2)$; the respective witness collection thus corresponds to a witness word of the form $v0^r a$ – where we suppose 0 to be the index of the empty segment (i.e., of the collection of empty sets) while a is the index of a nonempty (basic hom-) segment. A accepts the word $v0^r a$ according to our assumption. Due to the well-known pumping lemma, A would also accept a word $v0^s a$ where $s > r$; in the respective collection, $\mathcal{P}_2$ and $\mathcal{P}_5$ remain singletons but the condition $m_2 = \frac{5}{2} m_5$ is obviously violated – thus we have got a contradiction.

Remark. So it is clear that Bouziane's construction of the automaton (which uses further coding and other technical results) contains a serious mistake. The "finger" pointing at the mistake in the FOCS'98 paper would point at the following sentence in the proof of Lemma 3.4.: "If we put m_1 to be ... and we put m_2 to be ... then $m_1 = m_2$". There is no argument in the paper why this equality should hold. In fact, it could be contradicted by elaborating our counterexample but it would be a bit long and tedious. This false claim served for deriving a bound on the number of states of the constructed automaton – which in reality can have *infinitely* many states.

Acknowledgements. I thank André Arnold and Faron Moller for useful (email and personal) discussions.

References

[1] Bouziane Z.: A primitive recursive algorithm for the general Petri net reachability problem; in Proc. FOCS 1998, 130–136.

[2] Esparza J.: Decidability and Complexity of Petri Net Problems – an Introduction. Lectures on Petri Nets I: Basic Models. Advances in Petri Nets. G. Rozenberg and W. Reisig (Eds.). Lecture Notes in Computer Science 1491, 374–428 (1998).

[3] Lipton R.: The reachability problem is exponential-space-hard; Dept. Comp. Science Rep. 62, Yale Univ., New Haven, CT, 1976

[4] Mayr E.W.: An algorithm for the general Petri net reachability problem, in Proc. 13th Ann. ACM STOC, 1981, 238–246.

[5] Mayr E.W.: An algorithm for the general Petri net reachability problem, *SIAM J. Comput.* **13** (1984) 441–460.

[6] Mayr E.W., Meyer A.R.: The complexity of the word problems for commutative semigroups and polynomial ideals; *Advances in Mathematics* **46** (1982) 305–329.