Throughput Lower Bounds for Markovian Petri Nets: Transformation Techniques

Javier Campos, Beatriz Sánchez, and Manuel Silva
Dpto. de Ingeniería Eléctrica e Informática *
Centro Politécnico Superior, Universidad de Zaragoza
María de Luna 3, 50015 Zaragoza, SPAIN

Abstract

This paper addresses the computation of lower bounds for the steady-state throughput of stochastic Petri nets with immediate and exponentially distributed service times of transitions. We try to deeply bridge stochastic Petri net theory to untimed Petri net and queueing networks theories. Previous results for general service time distributions are improved for the case of Markovian nets by considering some pessimistic transformation rules operating locally on the net structure, its initial marking, and stochastic interpretation. Special interest have the obtained results for the case of live and bounded free choice nets and live marked graphs systems.

1 Introduction

The computation of lower bounds for the *throughput of transitions*, defined as the average number of firings per time unit, of stochastic Petri nets with immediate and exponentially distributed service times of transitions is considered in this paper. Previous results for general service time distributions $[2\star, 3\star, 4\star, 5\star]$[1], based only on the net structure, on the routing probabilities, and on the mean values of the service time of transitions, are improved. Throughput upper bounds for Markovian nets are studied in a companion paper [4].

The improvement of the throughput lower bound, for the particular case of exponential distributions, is based on a kit of *transformation rules*. The first, a *multistep preserving reduction rule*, allows to remove some places without changing the exact throughput of the net system model. The second group, *sequence and fork-join rules*, reduces some subnets to a transition, usually preserving the average firing time, but relaxing the probability distribution function (PDF) (e.g. relaxing from a series of exponentials to a single exponential of the same mean service time). The third group of transformation rules, *fork and join transition splitting*, produces some kind of "stochastic desynchronizations", in the net model. Transition splittings can be very pessimistic. Their interest becomes clear when combination with fork-join reduction rules is possible.

The mentioned kit of transformation rules allows to fully reduce any live and bounded 1-cut marked graph (i.e., such that there exists at least one transition belonging to all circuits). Live and bounded marked graphs (LBMG) can always be transformed into 1-cut marked graphs (1-cut MG) while making them slower by: (1) removing some circuits and changing the service time of some transitions or (2) changing the net structure in such a way that a new immediate transition belonging to all circuits is introduced.

The complete reduction of live and bounded free choice systems (LBFC) is impossible with our kit of reduction rules when we are looking for bounds. The reason is that choices lead to hyperexponential-like PDFs, that cannot be "stochastically smaller in mean" than the exponential services with the same mean value. Only if the reduced net has the topology of state machine, plus eventually some places self-loops around transitions, the reduction is totally satisfactory. In these cases a queueing network (QN) topology is obtained and classical product-form solutions hold, leading to polynomial time computation of the throughput for the transformed net system.

In general, the intuitive idea that decreasing the service time of a transition leads to a slower system is paradoxically wrong! Figure 1 shows a Markovian net system where increasing the service time s_2 of t_2, while $s_2 \in (0, 2)$, the throughput decreases. Therefore the idea of decreasing the throughput of a system by slowing down one subsystem (e.g., a transition, as in figure 1) does not hold in general, even if in practice this will be true for most cases. Therefore in general throughput approximations are computed. For the particular classes of LBFC nets and live MGs, a *performance monotonicity property* is shown to hold: *slowing down of a subsystem never allows better over-*

*This work was partially supported by the DEMON Esprit BRA 3148 and the PRONTIC 354/91.

[1]Citations labelled with $\star$ refer to the bibliography of the companion paper [4].

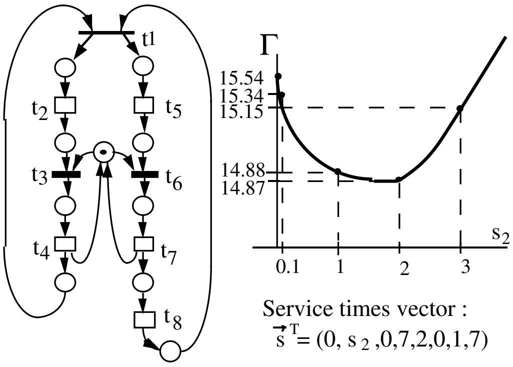

Figure 1: A first paradox: increasing the mean service time s_2 of t_2 on the Markovian net system leads to better overall throughput (Note: the underlying net is simple, mono-T-semiflow, and state machine decomposable).

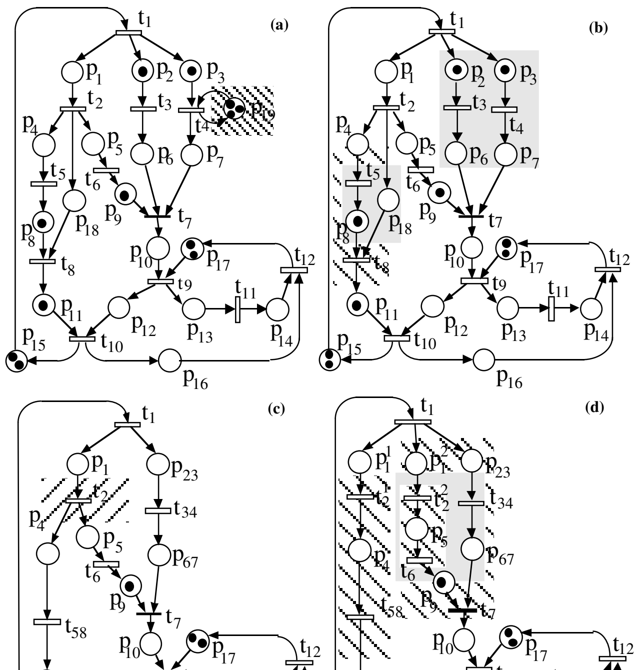

Figure 2: An LBMG and three first transformations preserving the throughput lower bound.

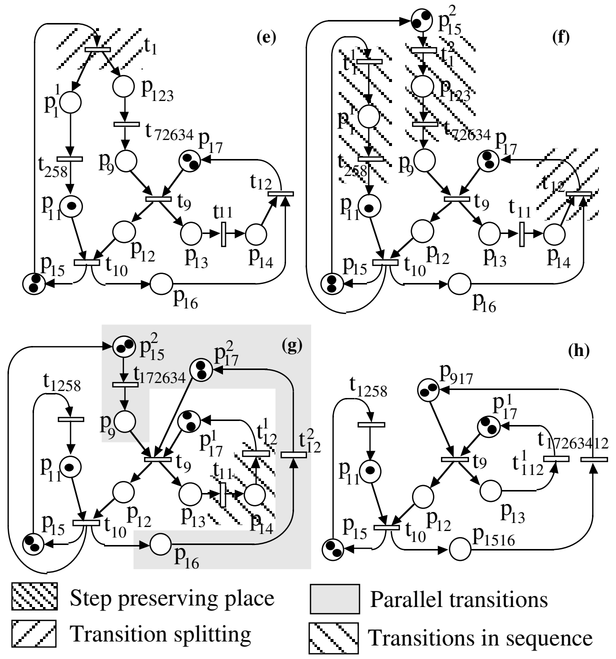

Figure 3: Additional transformations for the net in figure 2 preserving the throughput lower bound.

all throughput.

We assume that the reader is familiar with stochastic Petri net (PN) models. Some notations and considerations about the stochastic interpretation assumed in what follows can be seen in the companion paper [4].

The paper is organized as follows. In section 2 we recall the throughput lower bounds for stochastic PNs with general service time distributions (including both deterministic and stochastic timing) derived in previous works. In sections 3, 4, and 5, different transformation rules are introduced.

Figures 2 and 3 present a reduction process for an MG, using multistep preserving places, transitions in sequence, fork-join subnets, and fork-splitting transformation rules. Section 6 is devoted to state the completeness of the presented kit of transformation rules with respect to 1-cut MGs. Section 7 introduces some additional reduction rules making the all kit complete for the class of live (possibly unbounded) MGs. Conclusions are summarized in section 8.

2 Bounds for general service PDFs

In this section we recall [$3\star, 4\star, 5\star$] *insensitive* (i.e., for any PDF of services) throughput lower bounds for stochastic PNs. In what follows $\vec{\sigma}^*$ is the limit throughput vector and $\Gamma^{(j)} = 1/\vec{\sigma}^*_j$ the *mean interfiring time* of t_j (the inverse of its throughput).

The performance of a model with *infinite-server semantics* depends on the maximum degree of enabling of the transitions. For this reason we recall here two concepts of degree of enabling.

Definition 2.1 [5⋆] *Let $\langle \mathcal{N}, M_0 \rangle$ be a net system and t a given transition of $\mathcal{N}$. The enabling bound of t is $E(t) = \max\{k \mid \exists M \in R(\mathcal{N}, M_0) : M \geq kPRE[t]\}$. The liveness bound of t is $L(t) = \max\{k \mid \forall M' \in R(\mathcal{N}, M_0), \exists M \in R(\mathcal{N}, M') : M \geq kPRE[t]\}$.*

The enabling bound is a behavioural property, of complex computation. We recall its structural counterpart.

Definition 2.2 [5⋆] *Let $\langle \mathcal{N}, M_0 \rangle$ be a net system. The structural enabling bound of a given transition t of $\mathcal{N}$ is:*

$$SE(t) = \max\{k | M_0 + C \cdot \vec{\sigma} \geq kPRE[t]; \vec{\sigma} \geq 0\} \tag{LPP1}$$

The following result relates the three above presented concepts:

Theorem 2.1 [4⋆] *Let $\langle \mathcal{N}, M_0 \rangle$ be a net system.*
i) $\forall t \in T$: $SE(t) \geq E(t) \geq L(t)$.
ii) If $\langle \mathcal{N}, M_0 \rangle$ is reversible (i.e., M_0 is a home state), then $\forall t \in T$: $E(t) = L(t)$.
iii) If $\langle \mathcal{N}, M_0 \rangle$ is an LBFC system, then $\forall t \in T$: $SE(t) = E(t) = L(t)$.

Theorem 2.1.iii allows, for LBFC systems, to compute $L(t)$ in polynomial time, solving (LPP1). According to the next results, insensitive lower bounds on throughput for LBMGs and LBFC net systems can be computed in polynomial time.

Theorem 2.2 [3⋆, 4⋆] *For any LBFC net with a specification of the mean service times s_j for each $t_j \in T$ it is not possible to assign PDFs to the transition service times such that the mean interfiring time of transition t_i is greater than $\Gamma^{ub}_{(i)} = \sum_{j=1}^{m} v_j^{(i)} \frac{s_j}{SE(t_j)}$, independently of the topology of the net, where $v_j^{(i)}$ is the visit ratio to transition t_j, normalized for t_i (i.e., $v_i^{(i)} = 1$). For LBMGs, $\vec{v}^{(i)} = \mathbb{1}$ and the bound is reachable for some assignement of PDFs to the service time of transitions.*

The interpretation of the above result is as follows: a tight upper bound for the mean interfiring time of an LBMG (all its transitions have equal throughput) can be obtained putting in *series* the activity of all *transitions* (*stations* in QN terminology) and with a number of *servers* equal to its liveness bound.

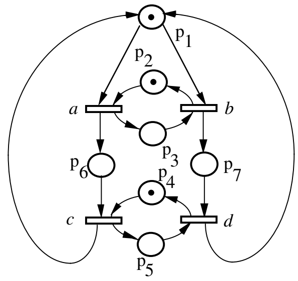

Figure 4: A second paradox: removing p_3 all liveness bounds go down to zero (i.e., the resulting net system has a total deadlock).

Liveness bounds of transitions of MG in figure 2.a are $L = (3,3,3,3,4,3,3,3,2,2,2,2)^T$. Assuming the following vector of service times: $\vec{s} = (1,1,10,10,6,14,0,2,1,1,1,1)^T$, the insensitive upper bound for the mean interfiring time is (theorem 2.2): $\Gamma^{ub} = 16.16$. This value can be reached for some PDFs. Nevertheless, if services are exponential the exact mean interfiring time of transitions is $\Gamma = 6.9665$.

In the following, information about the PDFs will allow us to improve our estimation of throughput lower bounds.

3 Multistep preserving places

In this section we are looking for a reduction rule that preserves all performance indices of a stochastic net system, for all possible stochastic interpretations. Thus we are looking for a reduction rule that, on the underlying net system, preserves the set of servers that can simultaneously work. According to this, transitions cannot be removed and our reduction rule must consider only places. More technically speaking, we are interested in those places whose elimination preserves the multisets of transitions simultaneously firable.

Places represent constraints to the firing of transitions of a net model. Therefore, removing a place usually leads (e.g., p_{18} in figure 2.a) but not always to an increase of the liveness bounds of transitions. Nevertheless, a counterintuitive fact is shown in figure 4. In some other cases the elimination of a place does not change the liveness bound of any transition (e.g., p_{19} in figure 2.a).

The elimination of *implicit places* [10⋆] preserves the firing sequences of the net system. Nevertheless concurrency and mutual exclusion properties are not preserved. Therefore a generalization of the implicit place concept is needed: *multistep preserving place*.

Let $\langle \widetilde{\mathcal{N}}, \widetilde{M_0} \rangle$ be the net system obtained by removing place p in $\langle \mathcal{N}, M_0 \rangle$.

Definition 3.1 *Let $\langle \mathcal{N}, M_0 \rangle$ be a net system. A multiset of transitions $\mathcal{S}$ is a multistep of $\langle \mathcal{N}, M_0 \rangle$ iff $\exists M \in R(\mathcal{N}, M_0)$ s. t. $M \geq PRE \cdot \vec{\mathcal{S}}$, where $\vec{\mathcal{S}}$ is the characteristic vector of the multiset.*

Definition 3.2 [5] *The place p is multistep preserving in $\langle \mathcal{N}, M_0 \rangle$ iff its elimination preserves the multisteps.*

From the above definition the next result can be stated:

Theorem 3.1 *Let p be a multistep preserving place in the live and structurally bounded system $\langle \mathcal{N}, M_0 \rangle$. Removing p preserves the following properties:*
1. Structural boundedness.
2. The enabling and liveness bounds, thus liveness and all the performance indices (throughput, queue lengths,...) for any stochastic interpretation (Markovian or not).
3. The number of reachable markings, thus the size of the embedded Markov chain of the Markovian PN.

Basically, removing multistep preserving places allows the application of other transformation rules as those presented in the next paragraphs.

Because $SE(t) \geq E(t) \geq L(t)$ (theorem 2.1), the following characterization can be done using convex geometry/linear programming theory. The result derives in a straighforward way from [5] and it is close to results in [10$\star$].

Let $C(p)$ be the row associated to p in C, and $\widetilde{C}$ the incidence matrix obtained by removing $C(p)$ in C.

Theorem 3.2 [5] *Assuming $\mathcal{N}$ is structurally live and structurally bounded, a sufficient condition for p being multistep preserving in $\langle \mathcal{N}, M_0 \rangle$ is:*
i) $\exists Y \geq 0$ s. t. $Y^T \cdot \widetilde{C} = C(p)$.
ii) $M_0(p) = Y^T \cdot \widetilde{M}_0 + \delta$, where $\delta = \max\{PRE[p] \cdot \vec{\mathcal{S}} | M = M_0 + C \cdot \vec{\mathcal{S}}; M \geq PRE \cdot \vec{\mathcal{S}}; M, \vec{\sigma}, \vec{\mathcal{S}} \geq 0$

Theorem 3.2 says that the places we will be able to remove are *marking structurally implicit* [10$\star$] (condition i), provided with a certain initial marking (condition ii). Places p_{18} and p_{19} in figure 2.a accomplish condition i.

The above results allow to state a reduction rule as follows. Note that there are no conditions/changes at stochastic level.

$\mathbf{R}_0$: Reduction Rule for Multistep Preserving Places:
Structural conditions: $\exists Y \geq 0$ s. t. $Y^T \cdot \widetilde{C} = C(p)$.
Marking conditions: $M_0(p) \geq Y^T \cdot \widetilde{M}_0 + \delta$.
Stochastic conditions: None.
Structural changes: $\widetilde{P} = P \backslash \{p\}$; $\widetilde{F} = F \backslash \{(t,p),(p,t)\}$
Marking changes: None.
Stochastic changes: None.

Place p_{19} in figure 2.a is multistep preserving ($Y = 0$, $\delta = 2$).

4 Rules almost-preserving local mean traversing time

In this section, we introduce two local reduction rules that almost preserve the mean traversing time of the subnet being reduced. This means that the rules presented here preserve the *mean traversing time of the reduced subnet by a single token*, in the simplest cases (we define the mean traversing time of one subnet by a token as the mean time elapsed since the token enters in the subnet until it goes out). In the most general cases, the mean traversing time of the subnet by a token can be increased by the rules. In all cases, the mean interfiring time of the whole net is increased. Thus a lower bound of throughput can be computed.

In particular, *sequences of transitions* (with exponentially distributed service times) and *fork-join subnets* (including parallel transitions with exponential service times) are reduced to single transitions with exponentially distributed service time. Although they preserve the mean traversing time of the reduced subnet, the random traversing time of that subnet is increased in the sense of a *stochastic ordering relation.* Technical details concerning this ordering relation are collected in section 4.1. The reduction rules are formally presented in section 4.2, together with the effect whose application has in the global throughput for the case of LBFC nets (and in particular for LBMGs).

4.1 Stochastic order relation

Let us suppose that two non-negative random variables (r.v.), X and Y, are such that Y gives *more weight to the extreme values* than X does; in other words, *Y has more variability than X*. This fact, which can be formally stated saying that $E[g(X)] \leq E[g(Y)]$ for all nondecreasing convex functions g, can be shown to be equivalent [1] to the next order relation between r.v.:

Definition 4.1 [1] *Let X and Y be non-negative r.v. with PDFs F and G, respectively. X is said to be stochastically smaller in mean than Y, written $X \prec Y$, iff for all $x \geq 0$, $\int_x^\infty \overline{F}(t)dt \leq \int_x^\infty \overline{G}(t)dt$, where $\overline{F} = 1 - F$ and $\overline{G} = 1 - G$.*

In particular, if X and Y have the same mean value, $X \prec Y$ is one way of guaranteeing that $Var(X) \leq Var(Y)$, because $g(x) = x^2$ is convex and $E[X] = E[Y]$.

Now, we present a class of r.v. that are stochastically smaller in mean (i.e., they have less variability) than exponential r.v., provided that they have the same mean values.

Definition 4.2 [1] *A non-negative r.v. X with PDF F is said to be new better than used in expectation (NBUE) iff its mean value $E[X]$ is finite and $\int_a^{\infty} \overline{F}(t)dt \leq E[X]\overline{F}(a)$ for all $a \geq 0$ such that $\overline{F}(a) > 0$.*

For the exponential PDF, the inequality of definition 4.2 holds as strict equality, so this type of PDF can be regarded as extreme in the class of NBUE. The next theorem referes to this fact:

Theorem 4.1 [1] *If X is NBUE and Y is exponential such that $E[X] = E[Y]$, then $X \prec Y$.*

In the sequel, we recall that sum and "max" operators over exponentially distributed r.v. give NBUE PDFs, that can be upper bounded, in ($\prec$)–relation sense, with exponential r.v. with the same mean (according to theorem 4.1).

Theorem 4.2 [1] *If $X_1, \ldots, X_n$ are independent and exponential r.v., then $X_1 + \cdots + X_n$ is NBUE distributed.*

Then, from theorems 4.1 and 4.2, the next result follows:

Corollary 4.1 *If $X_1, \ldots, X_n, Y$ are independent and exponential r.v. with $E[Y] = E[X_1] + \cdots + E[X_n]$, then $X_1 + \cdots + X_n \prec Y$.*

Analogous results hold for the "max" operator:

Theorem 4.3 [1] *If $X_1, \ldots, X_n$ are independent and exponential r.v., then $\max\{X_1, \ldots, X_n\}$ is NBUE distributed.*

Corollary 4.2 *If $X_1, \ldots, X_n, Y$ are independent and exponential r.v. with $E[Y] = E[\max\{X_1, \ldots, X_n\}]$, then $\max\{X_1, \ldots, X_n\} \prec Y$.*

We remark that, if $X_1, \ldots, X_n$ are independent and exponentially distributed (with $s_i = E[X_i]$), the mean value of the r.v. $\max\{X_1, \ldots, X_n\}$ can be easily computed as [1]:

$$E[\max\{X_1, \ldots, X_n\}] = h(s_1, \ldots, s_n) = \sum_{i=1}^{n} s_i - \\ - \sum_{\substack{i,j=1 \\ i<j}}^{n} \left(\frac{1}{s_i} + \frac{1}{s_j}\right)^{-1} + \sum_{\substack{i,j,k=1 \\ i<j<k}}^{n} \left(\frac{1}{s_i} + \frac{1}{s_j} + \frac{1}{s_k}\right)^{-1} + \\ + \cdots + (-1)^{n+1}\left(\frac{1}{s_1} + \cdots + \frac{1}{s_n}\right)^{-1} \tag{1}$$

In the next subsection, corollaries 4.1 and 4.2 justify the use of two reduction rules of sequential and parallel transitions for the computation of throughput lower bounds of LBFC nets (and in particular for LBMGs): they are based on the substitution of the sum and the maximum (respectively) of exponential r.v. for other exponential variable with equal mean value.

4.2 The reduction rules

Let us formally introduce the reduction rules for transitions in sequence and for parallel transitions. In the simplest cases, the mean traversing time of the reduced subnets are preserved and the throughput of the whole net decreases. The rules for the most general cases reduce the throughput of the net and also increase the mean traversing time of the subnets.

4.2.1 Transitions in sequence

Intuitively, this transformation makes indivisible the service time of two transitions representing elementary actions which always occur one after the other and leads to no side condition.

Definition 4.3 *Two transitions t_1 and t_2 are in sequence iff $t_1^{\bullet} = \{p_{12}\} = {}^{\bullet}t_2$.*

Note that more general cases with several places between transitions t_1 and t_2 can be reduced to the elementary one defined above, after the elimination of multistep preserving places (reduction rule presented in section 3).

Given two transitions in sequence, t_1 and t_2, with exponential services and means s_1 and s_2, a direct application of corollary 4.1 suggests to substitute both transitions and the connecting place p_{12} for a single transition t_{12} sum of t_1 and t_2 with exponential service and mean $s_1 + s_2$.

$\mathbf{R}_1$: Reduction Rule for Sequences:
Structural conditions: $\langle \mathcal{N}, M_0 \rangle$ contains two transitions in sequence, t_1 and t_2, connected through p_{12}.
Marking conditions: None.
Stochastic conditions: The r.v. X_1 and X_2 associated to the transitions t_1 and t_2 must be exponential (it holds also for NBUE PDFs).
Structural changes: $\widetilde{P} = P \setminus \{p_{12}\}$; $\widetilde{T} = T \setminus \{t_1, t_2\} \cup \{t_{12}\}$, where $t_{12} \notin T$; $\widetilde{F} = F \setminus \{(t_1, p_{12}), (p_{12}, t_2)\} \cup \{(p, t_{12}) | p \in {}^{\bullet}t_1\} \cup \{(t_{12}, p) | p \in t_2^{\bullet}\}$.
Marking changes: $\widetilde{M_0}(p) = M_0(p)$ if $p \notin {}^{\bullet}t_1$, $\widetilde{M_0}(p) = M_0(p) + M_0(p_{12})$ if $p \in {}^{\bullet}t_1$.
Stochastic changes: The r.v. X_{12} associated to t_{12} has an exponential PDF with mean $E[X_{12}] = E[X_1] + E[X_2]$. The (exponential) r.v. associated to other transitions remain unchanged.

According to corollary 4.1, the random service time of the new transition t_{12} is stochastically greater in mean than the random traversing time of the subnet consisting of t_1, p_{12}, and t_2, while the mean value of the traversing time is preserved. Concerning the whole throughput, the previous reduction usually produces a slower one. This fact is guaranteed if LBFC systems are considered.

Theorem 4.4 *Let $\langle \mathcal{N}, M_0 \rangle$ be a net system and $\langle \widetilde{\mathcal{N}}, \widetilde{M_0} \rangle$ that obtained applying R_1. The following statements hold:*

1. Liveness and boundedness are preserved.

2. The size of the state space is reduced.

3. If $\mathcal{N}$ is LBFC, then: (a) the reduced net $\widehat{\mathcal{N}}$ is also LBFC and (b) the throughput of any transition (different from t_1 and t_2) of the reduced net is less than or equal to the original one.

Proof: Statements 1, 2, and 3.a are obvious. Concerning the throughput, it must be pointed out that:
1. Either there exists a *fork-join subnet* including the reduced transitions whose mean traversing time is augmented by the reduction rule or the mean interfiring time of transitions in the whole net is preserved.
2. In a LBFC net, if the mean traversing time of a subnet is increased, then the throughput of the whole net decreases.

In order to prove claim 1, consider two transitions t_a and t_b of the net such that $|t_a^\bullet| > 1$, $|^\bullet t_b| > 1$, and such that there exists a directed path from t_a to t_1 and other directed path from t_2 to t_b (where t_1 and t_2 are the transitions in sequence being reduced). If such transitions t_a and t_b do not exist, then the considered net is a state machine, thus it is isomorphic to a product-form closed monoclass queueing network with infinite servers in all stations (*delay stations*), and the result stated in the theorem is true (the throughput does not change after the reduction). Assume that t_a and t_b exist. Then the mean traversing time (by a single token) of the subnet generated by the paths connecting t_a with t_b is less than or equal to the mean traversing time (by a single token) of the same subnet after the reduction (notice that the subnet includes transitions t_1 and t_2). The proof of this claim, that uses the results presented in section 4.1, can be found in [7], in the framework of *PERT models* with exponential timing. It is based on the fact that the mean value of the maximum among sums of NBUE r.v. is less than or equal to the mean value of the maximum among the corresponding sums of exponential r.v. with the same mean than the original NBUE PDFs.

In order to prove claim 2, we recall that LBFC nets can be decomposed into several state machines (P-components) connected by means of synchronization transitions [3]. Moreover, from the definition of FC nets, if p_a and p_b are input places to a synchronization transition t, then t is the unique output transition of p_a and p_b. In other words, once a synchronization transition has been enabled in an FC net, its firing is unavoidable. Then, if the mean (by a single token) time of a subnet increases after the reduction rule, the effect can be a greater amount of waiting time at synchronization transitions, but never the increasing of the throughput of transitions. ∎

As an example, transitions t_2^1 and t_{58} in the net depicted in figure 2.d are in sequence, and can be reduced to a single transition t_{258}.

The net in figure 1 is simple, mono-T-semiflow and state machine descomposable. This leads to the conclusion that the decrease of the throughput of the transformed net is not guaranteed for such net subclasses that are so close to FC nets in the net subclasses hierarchy.

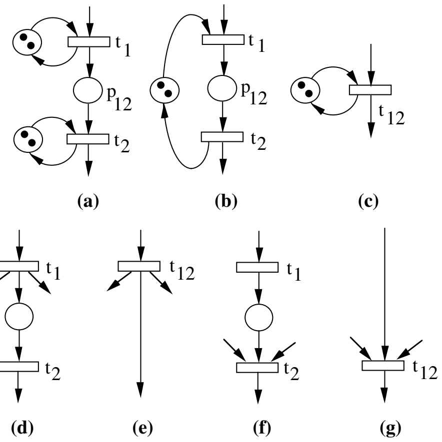

Figure 5: Generalizations of the reduction of sequences.

Some generalizations of the reduction rule presented above can be considered:
1. When transitions in sequence have their liveness bounds reduced by means of self-loop places (see figure 5.a). In these cases, a first pessimistic transformation can be made substituting the self-loop places by a single place connecting the two transitions in sequence, like is depicted in figure 5.b. This transformation mantains the liveness bound of both transitions but reduces the marking bound of place between them. After that, the reduction rule presented above can be aplied to the transitions in sequence, leading to the subnet of figure 5.c. We remark that this generalization of the rule does not preserve the mean traversing time of the reduced subnet: the mean traversing time is increased, therefore the reduced net is slower than the original. From a QN prespective the reduction consist in removing the queue represented by p_{12} and fusing stations t_1 and t_2 into a single one with mean service time $s_1 + s_2$.

2. Literature on reduction/transformation techniques for the qualitative analysis of PNs (e.g., [2, 8]) gives ideas on possible rules for stochastic PNs. Our stochastic reductions are based on the substitution of some subnets for exponentially distributed transitions, thus increasing the r.v. in the ($\prec$)–relation sense, and increasing the coefficient of variation to 1 (see section 4.1). Therefore, rules including choices and multiple attributions must not be considered, since they lead to hyperexponential PDFs (with coefficients of variation greater than 1) that cannot be stochastically bounded with exponentials. However, some classical reduction rules (without choices and multiple attributions) like those depicted in figure 5.d and 5.f can be considered. The transitions t_1 and t_2 in figures 5.d and 5.f can be reduced to the single transition t_{12} in figures 5.e and 5.g, respectively, with exponential service time, and mean $s_{12} = s_1 + s_2$. The throughput of the obtained net is, in general, less than the original.

4.2.2 Fork-join subnets

In this section, corollary 4.2 is used in order to reduce a *fork-join subnet* to a single transition. The most simple case of fork-join subnet that can be considered is depicted in figure 6.a. In this case, if tran-

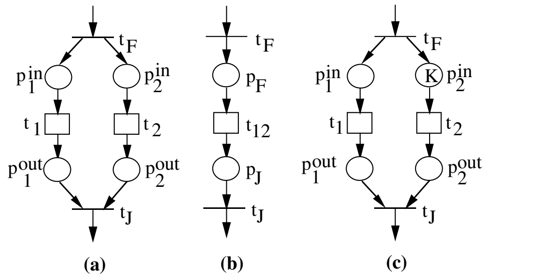

Figure 6: Elementary fork-joins and their reduction.

sitions t_1 and t_2 have exponential services X_1 and X_2 with means s_1 and s_2, they are reduced to a single transition (figure 6.b) with exponential service time and mean $E[\max\{X_1, X_2\}] = h(s_1, s_2)$, where h is the function defined in equation (1). Therefore, the mean traversing time of the reduced subnet by a single token has been preserved. However, by corollary 4.2 the random service time of the new transition t_{12} is stochastically greater in mean than the random traversing time of the reduced subnet (which is the maximum between two exponential r.v.). Concerning the throughput of the whole net, similar arguments to those in theorem 4.4 can be used to assure that the throughput of the net derived after reduction is smaller than the one of the original net. A trivial extension can be applied if the fork-join subnet includes more than two transitions in parallel, using corollary 4.2.

In a more general situation, depicted in figure 6.c, several tokens are trapped in the fork-join, and the mean traversing time of the subnet is less than the one of the subnet of figure 6.a. A pessimistic assumption in this case consists in supposing that the K trapped tokens are always in place p_2^{in} when transition t_F is fired. Then the mean traversing time of the subnet can be upper bounded by $E[\max\{X_1, \min\{\overbrace{X_2, \ldots, X_2}^{K+1}\}\}] = h(s_1, \frac{s_2}{K+1})$, where h is once more the function defined in equation (1). And now, since $\min\{X_2, \ldots, X_2\}$ is exponentially distributed, the subnet can be reduced to a single transition with exponential service time and mean equal to $h(s_1, s_2/(K+1))$, increasing the random traversing time in the ($\prec$)–relation sense (by corollary 4.2). The reader can notice that, before the application of previous reduction, "non-trapped tokens" must be removed from the inside of the fork-join and added to the input places of transition t_F.

In figure 7.a an example of the most general case of fork-join that we consider is depicted. In this case, some tokens can be trapped in the fork-join, and also self-loop places can limit the liveness bound of involved transitions. We present now the formal definition of the reduction rule, for these general cases (in fact, more than two transitions in parallel can be involved in the fork-join).

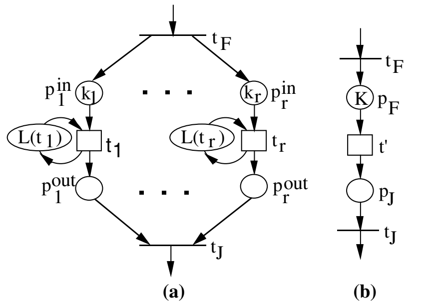

Figure 7: A more general fork-join.

R$_2$: Reduction Rule for Fork-Join Subnets:

Structural conditions: $\langle \mathcal{N}, M_0 \rangle$ contains a set $\overline{T} = \{t_1, \ldots, t_r\}$ of parallel transitions included between a fork transition, t_F, and a join transition, t_J.

Marking conditions: None.

Stochastic conditions: The r.v. $X_1, \ldots, X_r$ associated to the transitions in $\overline{T}$ must be exponential, with means $s_1, \ldots, s_r$.

Structural changes: Let $\overline{P} = {}^\bullet(\overline{T}) \cup (\overline{T})^\bullet$. Then: $\widetilde{P} = P \setminus \overline{P} \cup \{p_F, p_J\}$, where $p_F, p_J \notin P$; $\widetilde{T} = T \setminus \overline{T} \cup \{t'\}$, where $t' \notin T$; $\widetilde{F} = F \cap ((\widetilde{P} \times \widetilde{T}) \cup (\widetilde{T} \times \widetilde{P})) \cup \{(t_F, p_F), (p_F, t'), (t', p_J), (p_J, t_J)\}$.

Marking changes: Let be $k_i = M_0(p_i^{in}) + M_0(p_i^{out})$, where $p_i^{in} = t_F^\bullet \cap {}^\bullet t_i$, and $p_i^{out} = t_i^\bullet \cap {}^\bullet t_J$, $i = 1, \ldots, r$. Let be $k = \min\{k_i | i = 1, \ldots, r\}$. Then, $\widetilde{M}_0(p) = M_0(p)$ if $p \in P \setminus \overline{P}$, $\widetilde{M}_0(p) = k$ if $p = p_F$, $\widetilde{M}_0(p) = 0$ if $p = p_J$.

Stochastic changes: The r.v. X' associated to the transition t' has an exponential PDF with mean $E[X'] = h\left(\frac{s_1}{\min\{L(t_1), k_1 - k + 1\}}, \ldots, \frac{s_r}{\min\{L(t_r), k_r - k + 1\}}\right)$, where $k_1, \ldots, k_r, k$ are the constants defined above, and $L(t_i)$ is the liveness bound of t_i, $i = 1, \ldots, r$. The (exponential) r.v. associated to other transitions remain unchanged.

Note that if t_i are self-loop free then $k_i - k = L(t_i) - L(t_F), i = 1, \ldots, r$.

According to corollary 4.2, the following result can be derived, with similar arguments to theorem 4.4. In fact it follows easily for safe nets, while in the case of non-safe nets it is necessary to check that the mean traversing time of k tokens through the fork-join (in which some additional tokens can be trapped) is less than the mean traversing time of the same k tokens through the new transition resulting from the reduction. This computation can be made straightforward

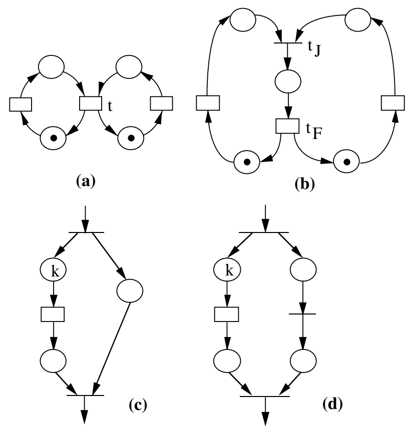

Figure 8: Particular cases of fork-joins.

using order statistics of exponential r.v. (see, e.g., [6]). We omit it by the lack of space.

Theorem 4.5 *Let $\langle \mathcal{N}, M_0\rangle$ be a net system and $\langle \widetilde{\mathcal{N}}, \widetilde{M_0}\rangle$ that obtained applying R_2. The following statements hold:*

1. Liveness and boundedness are preserved.

2. The size of the state space is reduced.

3. If $\mathcal{N}$ is LBFC, then: (a) the reduced net $\widehat{\mathcal{N}}$ is also LBFC and (b) the throughput of any transition (different from those in $\overline{T}$) of the reduced net is less than or equal to the original one.

As an example, let us consider again the net in figure 2. Transitions t_3 and t_4 in figure 2.b belong to a fork-join, thus they can be reduced to a single transition t_{34} applying the above rule (figure 2.c).

The reader can notice that some particular schemes, seemingly different to the previous general one, can be easily transformed into fork-join subnets. As an example consider the net in figure 8.a. For this net, the fork and join transitions are the same: t. They can be made visible, by splitting t into two different transitions t_F and t_J with a connecting place (see figure 8.b). After that, the fork-join reduction rule can be applied.

Another particular scheme is depicted in figure 8.c. In this case, the place p can be split into two places connected through an immediate transition (see figure 8.d) and the fork-join reduction rule can also be applied.

5 Splitting transitions

Multistep preserving places, sequence transitions, and fork-join subnets are reduction rules leading to net systems with smaller state (marking) space. Nevertheless, they do not allow the reduction of net systems like that in figure 2.c.

In this section a new transformation rule is introduced. It is not a reduction rule in the sense that the state space of the transformed net system tends to be larger than the original one. From a performance point of view, it does not preserve the mean traversing time of the subnet involved in the transformation, because it replicates the transition to split. So, where are the advantages of this new rule ? The answer is that it allows to proceed further in the reduction process using again the previous rules.

The figure 2.c and 2.d show how the transition t_2 is split into two transitions t_2^1, t_2^2. This is called a *fork-splitting*, because the split transition is a fork one (i.e., $\mid t^\bullet \mid > 1$).

After that splitting for the net in figure 2.d, $\vec{\sigma}^* = 0.1350$ and $\vec{\sigma}^{lb} = 0.0718$, and the state space has grown from 1011 to 3066 markings. Nevertheless, the application of sequence and fork-join rules to the net of figure 2.d leads to the system in figure 3.e. In table 1, it is shown that the error of the insensitive bound computed with the system in figure 2.c is 46% and in 3.e only 36%.

Definition 5.1 *Let $\mathcal{N} = \langle P, T, F\rangle$ be a net, $t_F \in T$ is a* 1-fork transition *iff:*
i) t_F is a fork transition (i.e. $\mid t_F^\bullet \mid = q > 1$)
ii) ${}^\bullet t_F = \{p_F\}$ and $\mid p_F^\bullet \mid = \mid {}^\bullet p_F \mid = 1$

Let t_F be a fork transition, $t_F^\bullet = \{p_{i+1}^j \mid j : 1, \ldots, q\}$ where $q = \mid t_F^\bullet \mid$, and ${}^\bullet p_F = \hat{t}$. The r.v. associated to t_F is X_F, exponentially distributed with mean s_F.

R$_3$: Transformation Rule for Fork Splitting:
Structural conditions: $\mathcal{N}$ has a 1-fork transition t_F.
Marking conditions: None.
Stochastic conditions: None.
Structural changes: $\widetilde{P} = P \setminus \{p_F\} \cup \{p_F^j \mid j = 1, \ldots, q\}$, $p_F^j \notin P$; $\widetilde{T} = T \setminus \{t_F\} \cup \{t_F^j \mid j = 1, \ldots, q\})$, $t_F^j \notin T$; $\widetilde{F} = F \setminus \{(p_F, t_F), (t_F, p_{i+1}^j) \mid j = 1, \ldots, q\} \cup \{(\hat{t}, p_F^j), (p_F^j, t_F^j), (t_F^j, p_{i+1}^j) \mid j = 1, \ldots, q\}$.
Marking changes: $\widetilde{M_0}(p) = M_0(p)$ if $p \in P \setminus \{p_F\}$; $\widetilde{M_0}(p_F^j) = M_0(p_F) \forall j = 1, \ldots, q$.
Stochastic changes: The new transitions, denoted t_F^j, $j = 1, \ldots, q$ have associated r.v. X_F^j, $j = 1, \ldots, q$ with the same PDF than X_F.

With similar arguments to theorem 4.4, the following theorem can be proved:

Theorem 5.1 *Let $\langle \mathcal{N}, M_0\rangle$ a net system and $\langle \widetilde{\mathcal{N}}, \widetilde{M_0}\rangle$ the obtained fork-splitting transition t_F.*

1. $\widetilde{\mathcal{N}}$ is structurally live and structurally bounded iff $\mathcal{N}$ is (i.e. structural liveness and structural boundedness are preserved).

2. The state space of $\langle \widetilde{\mathcal{N}}, \widetilde{M_0}\rangle$ is never smaller than that of $\langle \mathcal{N}, M_0\rangle$.

3. If $\mathcal{N}$ is LBFC then: (a) $\widetilde{\mathcal{N}}$ is also LBFC and (b) the throughput of any transition of the reduced net is *less than or equal to the original one (equality holds only when the splitted transition is immediate).*

For join-splitting, denoted $\mathbf{R}_4$: **Join-Splitting**, formal definition (the reverse of fork-splitting) and transformation rules are not included here, by the lack of space. Finally, theorem 5.1 equally holds for join-splitting.

When we compute a throughput bound using fork-splitting or its reverse join-splitting, we will not obtain the same insensitive bound. These two possibilities and the choice of the transition to split, give us several heuristics for the transformation process, and also we get differents bounds. Some of these heuristics could be: (1) use only fork-splitting; (2) use only join-splitting; (3) choose among the candidate transitions to be split, the one with minimun mean service time and apply R_3 or R_4 according to the nature of the transition.

The quantitative results of the transformation process illustrated in figures 2 and 3 are shown in table 1. $|Tang|$ represents the cardinal of the tangible reachable markings set. Γ_{exact} correspond to the exact mean interfiring time of the net (in fact, of any transition). Γ^{ub}_{insens} is an upper bound for the mean interfiring time of transitions of the MG (computed using theorem 2.2). The relative error in the last column of each row in the table is computed between the exact throughput of the net in figure 2.a and the insensitive throughput lower bound of the net corresponding to that row ($1/\Gamma^{ub}_{insens}$). The estimated bound has been improved from 16.16 to 10.49 (56% error to 33% error). An additional improvement (15% error) is presented in the next section.

Table 1

Net	$\|Tang\|$	Γ_{exact}	Γ^{ub}_{insens}	Error
a (Fig.2)	8698	6.9665	16.16	56%
b (Fig.2)	8698	6.9665	16.16	56%
c (Fig.2)	1011	7.0414	12.91	46%
d (Fig.2)	3066	7.0421	13.91	49%
e (Fig.3)	187	7.1893	10.91	36%
f (Fig.3)	540	7.2225	11.91	41%
g (Fig.3)	336	7.2259	12.91	46%
h (Fig.3)	48	7.5847	10.49	33%

6 Complete reduction for 1-cut MGs

In the following we prove that the previous kit of transformation rules constitute a complete system to reduce any 1-cut MG to an *elementary net* containing a single transition and a single place. The subclass of 1-cut MGs represent a generalization of the classical PERT model, in such a way that cyclic behaviours can be modeled, as well as many different classes of non-shared resources for the realization of activities (tokens at places of the net). We impose the restriction that every cycle of the MG cross a common transition. This subclass is denoted 1-cut MGs, recalling the so-called cut set space of a graph (cut-set of a connected graph is a set of edges whose removal would disconnect the graph).

Definition 6.1 *Let $\mathcal{N} = \langle P,T,F\rangle$ be an MG. $\mathcal{N}$ is a* 1-cut MG *iff $\bigcap_{Y\in\{P-semiflows\}} ||Y^T \cdot PRE|| \neq \emptyset$, where $||Y^T \cdot PRE|| = \{t \in T | Y^T \cdot PRE > 0\}$ is the support of $Y^T \cdot PRE$.*

In order to prove that one set of reduction rules is complete we have to do three things: first, we construct a function associated with the net, and show that after applying anyone of the reduction rules, the value of the function decreases. Second, we prove that the elementary net has a minimum value for this function. And, third we prove that we always can apply some reduction rule until we get the elementary net.

Definition 6.2 *Let $\mathcal{N} = \langle P,T,F\rangle$ be a 1-cut MG, with $t \in \bigcap_Y ||Y^T \cdot PRE||$. The spread of $\mathcal{N}$, denoted by $SP(\mathcal{N},t)$, is the following pair of numbers:*

$$SP(\mathcal{N},t) = (|P|, \textstyle\sum_{l=1}^{m} d(t_l,t)\cdot(|t_l^\bullet|-1))$$

where $d(t_l,t)$ is the number of places between t and t_l in the minimum path which contains both.

In the sequel $\sum_{l=1}^{m} d(t_l,t)\cdot(|t_l^\bullet|-1)$ is denoted by $D(T)$.

Property 6.1 *$SP(\mathcal{N},t) = (1,0)$ for the elementary net $\mathcal{N}$ (with $P = \{p\}, T = \{t\}$, and $F = \{(p,t),(t,p)\}$).*

Definition 6.3 *Let $\mathcal{N} = \langle P,T,F\rangle$ and $\mathcal{N}' = \langle P',T',F'\rangle$ be MGs. We say that $SP(\mathcal{N},t) \leq SP(\mathcal{N}',t)$ iff:*
i) $|P| \leq |P'|$ and $D(T) = D(T')$ or
ii) $D(T) \leq D(T')$.

Theorem 6.1 *The rules $R_0,\ldots,R_4$ decrease the spread of N.*

Proof: *Multistep Preserving Places:* If we apply R_0 we eliminate the implicit places of the net. We have $\mathcal{N} = \langle P,T,F\rangle$ and SP($\mathcal{N}$,t) $= (|P|, D(T))$. After applying R_0 we have $\mathcal{N}' = \langle P',T,F\rangle$ with $SP(\mathcal{N}',t) = (|P|-k, D(T))$, where k is the number of multistep preserving places. Hence $SP(\mathcal{N}',t) \leq SP(\mathrm{N},t)$.
Transitions in sequence: If we apply R_1 between two transitions t_i and t_j we reduce the spread of the net. We have SP($\mathcal{N}$,t) $= (|P|, D(T))$. After R_1, $SP(\mathcal{N}',t) = (|P'|, D(T'))$ with: $D(T') = \sum_{\substack{l=1\\ l\neq i,j}}^{m} d(t_l,t)\cdot(|t_l^\bullet|-1)$ and $|P'| = |P|-1$.

If $d(t_i,t) = k$ then $d(t_j,t) = k+1$ because t_i and t_j are in sequence, then $|t_i^\bullet| = |{}^\bullet t_j| = 1$ and $|t_j^\bullet| = a$. So, it is easy to see that $D(T) = D(T') + a - 1$.
Fork-join subnets: When we apply R_2 to $t_{i_1}\ldots t_{i_k}$ we also reduce the spread of the net. $SP(\mathcal{N},t) = (|P|, D(T))$; after R_2 we have $SP(\mathcal{N}',t) = (|P'|, D(T'))$ with: $D(T') = D(T) - (|t_i^\bullet| - k - 1)\cdot d(t_i,t)$ and $|P'| = |P| - (2k-2)$ where t_i is the fork transition.
Splitting Transition: If we apply R_3 (or R_4 in the reverse) to t_j and t_i, it is not obvious that we reduce the spread of the net,

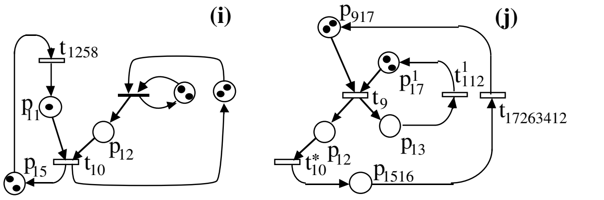

Figure 9: (j) Stochastic net system obtained from that in figure 3.h by removing the cycle with $p_{15} - t_{1258} - p_{11}$. The mean service time of $t_{10}^{\star}$ is computed with the net in (i).

while it is easy to see that we increase its size (number of places and transitions). If we make explicit the term of t_i and t_j in $D(T)$: $D(T) = \sum_{\substack{l=1 \\ l \neq i,j}}^{m} d(t_l, t) \cdot (|t_l^{\bullet}| - 1) + n \cdot (a-1) + (n+1) \cdot (k-1)$, where $d(t_j, t) = n$, $d(t_i, t) = n+1$, $|t_j^{\bullet}| = a$ and $|t_i^{\bullet}| = k$. After R_3 we have: $D(T') = \sum_{\substack{l=1 \\ l \neq j, i_j; 1 \leq j \leq k}}^{m} d(t_l, t) \cdot (|t_l^{\bullet}| - 1) + n \cdot (a + (k-1) - 1)$ and $|P'| = |P| + (k-1)$. It is easy to see that $D(T') \leq D(T)$ because $k \geq 2$.

So, we have proved that the value of function SP decreases when we apply any transformation rule to our net. ∎

Lemma 6.1 *Let $\mathcal{N}$ be a strongly connected MG, without structurally implicit places. If $\exists t$ s. t. $t \in \bigcap_Y ||Y^T \cdot PRE||$ (i.e., there exists at least one special transition included in all circuits), then $\{t^{\bullet}\} \cap \{{}^{\bullet}t'\} \neq \emptyset \Rightarrow |{}^{\bullet}t'| = 1$).*

Proof: Suppose that $|{}^{\bullet}t'| \geq 1$. Then $\{{}^{\bullet}t'\} \supseteq \{p, q\}$. As $\mathcal{N}$ is strongly connected, there exists a path from t to q ; if this path includes t' then there exists a cycle not including t. If the path does not include t' then p is a structurally implicit place. ∎

Theorem 6.2 *Let $\mathcal{N}$ be a 1-cut MG. The kit of transformation rules $\mathcal{C} = \{R_0, R_1, R_2, R_3\}$ reduces $\mathcal{N}$ to the elementary net.*

Proof: Let be $t \in \bigcap_Y ||Y^T \cdot PRE||$. In order to see that $\mathcal{C}$ is a complete set of transformation rules, we will prove that if we cannot use R_1, R_2 or R_0, then we can apply R_3.

Denote by $t_1, ..., t_l$ the transitions following t. As we cannot apply R_0, by lemma 6.1 we know that $|{}^{\bullet}t_i| = 1, i = 1, \ldots, l$. We also cannot apply R_1, so $l \geq 2$.

Then we know that: 1) $l \geq 2$; and 2) $|{}^{\bullet}t_i| = 1, i = 1, \ldots, l$.

Consider two cases. Case 1: $\exists i : 1 \leq i \leq l : |t_i^{\bullet}| \geq 2$. Then we can apply R_3. Case 2: $\forall i : 1 \leq i \leq l : |t_i^{\bullet}| = 1$. Denote t_i' the transitions following t_i for $i = 1, \ldots, l$ and $|t_i^{\bullet}| = p_i$.

If $\exists i, 1 \leq i \leq l$, s. t. $|{}^{\bullet}t_i'| = 1$, we could apply R_1 which contradicts the hypothesis. Then $|{}^{\bullet}t_i'| \geq 2, i = 1, \ldots, l$. We also know that $\forall i : 1 \leq i \leq l; p_i \in \{{}^{\bullet}t_j'\}; 1 \leq j \leq l$. If $i \neq j$ then we can apply R_2.

At this point we reach a contradiction. In order to prove it we define an orden relation between transitions in the following way: we say that t_i' is *lower* than t_j', denoted $t_i' \lhd t_j'$, if there exists a path from t_i' to t_j' which does not include t. It is easy to see that $\lhd$ is a strict partial order relation: (1) Irreflexive: If there exists a path from t_i' to t_i' not containing t, this means that there exists a cycle which does not include t. (2) Antisymmetric: By the same reason. (3) Transitive: Obvious.

There exists a minimal element in $t_1', ..., t_l'$. Let t_1' be this element, then $\forall m : 1 \leq m \leq l$ there exists a path from t_1' to t_m' that does not include t, but there does not exist a path from t_m' to t_1' not containing t, then $|{}^{\bullet}t_1'| = 1$, because in other case we have a cycle that does not include t. ∎

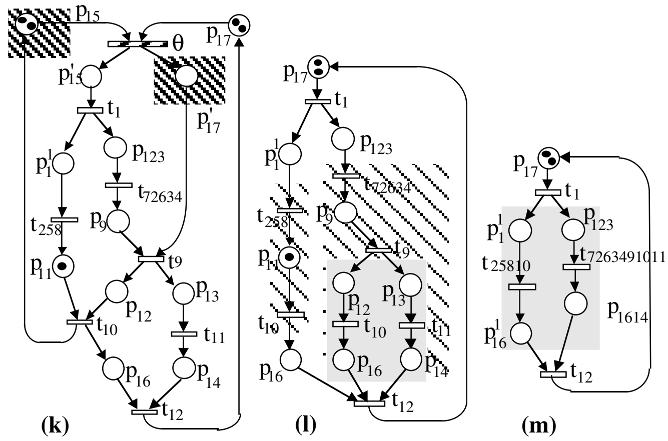

Figure 10: After the introduction of the immediate transition θ in the net in figure 3.e, a 1-cut MG is obtained (thus complete reduction is possible).

Therefore the kit of rules $\{R_0, R_1, R_2, R_3\}$ is complete for reducing the 1-cut MG subclass and the computing complexity is *polynomial* in the number of transitions of the net. In the same way, the complete reduction using R_4 instead R_3 is straightforward because the reverse of a 1-cut MG is also a 1-cut MG.

Theorem 6.3 *Let $\mathcal{N}$ be a 1-cut MG. The kit of transformation rules $\mathcal{C} = \{R_0, R_1, R_2, R_4\}$ reduces $\mathcal{N}$ to the elementary net.*

As a final remark, just point out that, even if theorems 6.2 and 6.3 say that from a logical point of view R_3 or R_4 can be indifferently used, the quantitative results may be not the same!

7 Additional transformation rules

The basic aim of this section is to show that all LBMGs can be reduced to 1-cut MGs, being possible to obtain, by transformations, a lower bound on throughput. The techniques will be presented in a generalized and formalized way in a forthcoming paper. Here we bound ourselves to the presentation of some basic ideas in an informal illustrative way.

Since the computation of the performance lower bound of live and unbounded MGs can be reduced to the computation of the performance bounds of several LBMG components [2⋆], a full reduction theory for live MGs is available.

Any LBMG (thus strongly connected) can be transformed in many ways into 1-cut MG by *removing some a few cycles*. For example, removing $p_{15} - t_{1258} - p_{11}$

in the MG of figure 3.h a 1-cut MG is obtained: all cycles go through t_9. A simple way of removing the cycle $p_{15} - t_{1258} - p_{11}$ is to substitute the service time of t_{10}, s_{10}, by $s^\star_{10} = s_{10} + s_{1258}$. The throughput of the resulting MG will be smaller and the previous reduction rules can be safely applied to compute a lower bound. Nevertheless, in general, $s^\star_{10} = s_{10} + s_{1258}$ may be too pessimistic because this means a *total sequentialization* between t_{10} and t_{1258}.

What about a more reasonable value for $s^\star_{10}$ leading to a slower system? One way of proceeding consists of making immediate the service time of t_9, t^1_{112}, and $t_{17263412}$. After the obvious reduction of immediate transitions (e.g, p_{13} and p^1_{17}, p_{16} and p^2_{17} are fused) the MG in figure 9.i is obtained. Removing the place self-loop on an immediate transition and fusing p_{12} with its predecessor, a particular case of the fork-join rule can be recognized and $s^\star_{10} = 3.72$. Therefore by removing the cycle $p_{15} - t_{1258} - p_{11}$ and substituting s_{10} by $s^\star_{10} = 3.72$, slower performance is obtained (figure 9.j). Now the MG is 1-cut. Once all computations have been done $\Gamma = 8.6101$ and $\Gamma^{ub} = 9.125$.

Table 2

Net	$\lvert Tang \rvert$	Γ_{exact}	Γ^{ub}_{insens}	Error
a (Fig.2)	8689	6.9665	16.16	56%
j (Fig.10)	18	8.9214	9.91	29%
Full red. of j	1	8.6101	9.125	23%
k (Fig.11)	85	7.9125	11.12	37%
l (Fig.11)	239	7.9150	11.45	39%
m (Fig.11)	14	7.9872	9.62	27%
Full red of m	1	8.23	8.23	15%

An alternative way of obtaining a 1-cut MG is *adding an immediate transition* such that all cycles go through it. The MG in figure 10.k has been obtained from the MG in figure 3.e by adding an immediate transition which synchronizes all circuits (P-invariants), preserving the liveness of the net. The introduction of the new transition clearly will decrease the performance of the original net. Thus, once again we are safely working to compute a lower bound on throughput. A subsequent reduction process is illustrated in figure 10. Once all computations have been done $\Gamma^{ub} = 8.23$. And its inverse is closer to the exact throughput of the original net (15% error, see table 2).

In general, neither technique is better than the other. The quality of the result depends on the structure, mean service times, and initial marking. The problem of which heuristic to apply at a certain moment of the transformation process, together with the formalization and generalization will be considered in detail in a forthcoming paper.

8 Conclusions

We have addressed the computation of *lower bounds for the throughput* of transitions in stochastic PN models (or the corresponding synchronized queueing networks). The class of PDFs associated to the firing of transitions has been defined as *exponential*, nevertheless for some rules (sequence and fork-join) any NBUE PDF can be considered. The improvement over the computation of the insensitive lower bounds (i.e. those based only on the net structure, the routing probabilities, and the mean service times) can be appreciable.

Technically speaking, a kit of transformation rules produces local pessimistic temporal behaviour, leading in general to an approximation for throughput. For the cases of LBFC systems (thus for LBMGs), lower bounds have been shown to be obtained. The introduced transformation rules can be classified from the performance perspective as follows: (1) Multistep preserving places rule: does not change the performance (for any stochastic interpretation). (2) Sequence and fork-join rules: preserve or almost-preserve the traversing time of the reduced subnet. (3) Transition splitting rules: destroy a fork or a join replicating the transition. They are interesting only if other transformation rules are applicable later, in particular the fork-join.

These rules allow to fully reduce 1-cut MGs, and have been proved useful in the reduction of the state (marking) space of many net systems. Finally, techniques to transform an MG into a 1-cut MG have been introduced in a simple illustrative way.

An extension of ideas presented in this paper concern the *approximate* analysis of performance. In this case, the presented rules plus other taking into account the reduction of choices are being considered. In order to derive "reasonable" approximations of true performance, both the first and second moments of PDFs are being kept in the transformation rules. The combination of lower bounds with upper bounds [4] and approximation techniques should be an important step to elaborate an "educated guess" of performance with reasonable computational cost.

References

[1] R. E. Barlow and F. Proschan. *Statistical Theory of Reliability and Life Testing.* Holt, New York, NY, 1975.

[2] G. Berthelot. Transformations and decompositions of nets. In W. Brauer, W. Reisig, and G. Rozenberg, ed., *Advances in PNs 1986*, vol. 254 of *LNCS*, pp. 359–376. Springer-Verlag, Berlin, 1987.

[3] E. Best. Structure theory of Petri nets: The free choice hiatus. In W. Brauer, W. Reisig, and G. Rozenberg, ed., *Advances in PNs 1986*, vol. 254 of *LNCS*, pp. 168–205. Springer-Verlag, Berlin, 1987.

[4] J. Campos and M. Silva. Throughput upper bounds for Markovian Petri nets: Embedded subnets and queueing networks. In *Proc. of PNPM'91*, Melbourne, Australia, Dec. 1991. IEEE-CS Press.

[5] J. M. Colom. *Análisis Estructural de Redes de Petri, Programación Lineal y Geometría Convexa.* PhD thesis, Dpto. Ingeniería Eléctrica e Informática, Univ. Zaragoza, Spain, June 1989.

[6] W. Feller. *An Introduction to Probability Theory and Its Applications.* John Wiley & Sons, New York, NY, 1966.

[7] J. Kamburowski. An upper bound on the expected completion time of PERT networks. *European Journal of Operational Research*, 21:206–212, 1985.

[8] M. Silva. *Las Redes de Petri en la Automática y la Informática.* Editorial AC, Madrid, 1985.