

Analysis of a Combined Queueing-Petri-Network World

Falko Bause, Heinz Beilner
Informatik IV, Universität Dortmund
Postfach 500 500, 4600 Dortmund 50
Germany

Topics:

- Theoretical developments in Timed and Stochastic Petri Nets, including high-level stochastic Nets
- Analysis methods for Timed and Stochastic Nets
- Fundamental advances in incorporating time into nets, while retaining the results of General Net theory

Abstract

System analysis is often needed with respect to both qualitative and quantitative aspects. In the last two decades, several model worlds have been developed that attempt to combine qualitative and quantitative aspects. Present emphasis is on Timed and Stochastic Petri Net models. One of the disadvantages of these model worlds lies with the difficulties when describing scheduling strategies with Petri Net elements.

This paper describes the QPN world, which combines Queueing Networks and Petri Nets, aiming at eliminating these disadvantages. The QPN world is a superset of Queueing Networks, Petri Nets and Timed and Stochastic Petri Nets. We also discuss the analysis of time-augmented Petri Nets in general. It is shown that several important qualitative features need not carry over to a Timed Petri Net, although they hold for the underlying "untimed" net. A solution to this problem is provided by showing that important qualitative features remain invariant, if the form of time integration is restricted by particular, easily testable conditions.

Contents

1	Introduction	1
2	Basic definitions and preliminary remarks	1
2.1	Petri Nets	1
2.2	Queueing Networks	3
2.3	Timed and Stochastic Petri Nets	5
3	The QPN world	7
3.1	Definition	7
3.2	Examples	8
3.3	Expressive Power	10
4	Combined Qualitative and Quantitative Analysis using QPNs	10
4.1	A general procedure for the analysis of time-augmented Petri Nets and associated difficulties	10
4.2	A solution strategy	11
4.2.1	Extended free choice QPNs	14
5	Conclusions	17

1 Introduction

Petri Nets and Queueing Networks are familiar model worlds for describing and analysing a variety of dynamic systems of the discrete-event type, such as computing systems (e.g. [27]) or manufacturing systems (e.g. [28]). Petri Nets are usually employed for assessing systemic correctness properties of such systems ("qualitative analysis") whereas the primary domain of Queueing Networks lies with assessing their performance properties ("quantitative analysis"). Taking an engineering point of view, model-based system analysis ought to be pursued with respect to all relevant system properties, in particular to both qualitative and quantitative ones. From this point of view, either one of the above model worlds is, in its pure form, incomplete and needs to be complemented by the other. To overcome these insufficiencies, both model worlds have (by themselves) been enhanced in various ways by additional modelling features, giving rise to new model worlds such as Timed and Stochastic Petri Nets (e.g. [21]) and Extended Queueing Networks (e.g. [29]). Those new model worlds are compared in [29] as regarding their expressive powers and their evaluation efficiencies. One important result of that study is, that certain specific deficiencies remain in both areas: Timed and Stochastic Petri Nets have, e.g., still difficulties in describing scheduling strategies in an appropriately compact manner; Extended Queueing Networks show, e.g., still deficits in describing sophisticated synchronisation aspects.

In this paper, a novel model world (the "QPN world") is introduced with the objective of combining the capabilities of Petri Nets and Queueing Networks and, consequently, eliminating the deficiencies of both (isolated) worlds. Moreover, the QPN world to be described forms a superset of (pure, Timed and Stochastic) Petri Nets and (standard and various Extended) Queueing Networks.

When enhancing pure Petri Nets with timing properties, a particularly dangerous aspect arises from the fact that qualitative model properties need not be invariant across an addition of time. This holds for almost all time-augmented Petri Nets (e.g. for Generalised Stochastic Petri Nets [1]) inclusive of the novel QPN world. To alleviate this situation, however, necessary and sufficient conditions will be provided for QPN models which guarantee an invariance of certain qualitative model properties across time-augmentation.

This article is structured as follows. Section 2 presents basic definitions. Section 3 provides the definition and some examples of the QPN world. In section 4, we discuss a general method for analysing time-augmented Petri Nets and show how this method can be used to analyse QPNs.

2 Basic definitions and preliminary remarks

2.1 Petri Nets

Definition 1 (Place/Transition Net) *A Place/Transition Net is a 5-tuple $PN = (P, T, I^-, I^+, M_0)$ where*

- P is a finite and non-empty set of places,
- T is a finite and non-empty set of transitions,
- $P \cap T = \emptyset$,
- $I^-, I^+ : P \times T \rightarrow \mathbb{N}_0$ are the backward and forward incidence functions,

- $M_0 : P \rightarrow \mathbb{N}_0$ is the initial marking.

Definition 2 (Basic notions of Place/Transition Nets) *Let*

$PN = (P, T, I^-, I^+, M_0)$ be a Place/Transition Net. $t\bullet := \{p \in P \mid I^+(p, t) > 0\}$, $\bullet t := \{p \in P \mid I^-(p, t) > 0\}$, $p\bullet := \{t \in T \mid I^-(p, t) > 0\}$, $\bullet p := \{t \in T \mid I^+(p, t) > 0\}$, and for $\tilde{T} \subseteq T, \tilde{P} \subseteq P : \tilde{T}\bullet := \bigcup_{t \in \tilde{T}} t\bullet$, $\bullet\tilde{T} := \bigcup_{t \in \tilde{T}} \bullet t$, $\tilde{P}\bullet := \bigcup_{p \in \tilde{P}} p\bullet$, $\bullet\tilde{P} := \bigcup_{p \in \tilde{P}} \bullet p$. $\tilde{P} \subseteq P$ is a deadlock, iff $\bullet\tilde{P} \subseteq \tilde{P}\bullet$. $\tilde{P} \subseteq P$ is a trap, iff $\tilde{P}\bullet \subseteq \bullet\tilde{P}$. A set $\tilde{P} \subseteq P$ is marked in a marking M , iff $\exists p \in \tilde{P} : M(p) > 0$; otherwise \tilde{P} is called unmarked or empty in marking M . A transition $t \in T$ is enabled in a marking M , denoted by $M[t >$, iff $M(p) \geq I^-(p, t), \forall p \in P$. An enabled transition $t \in T$ may fire. The firing of $t \in T$ yields a new marking given by $M'(p) = M(p) + (I^+(p, t) - I^-(p, t)), \forall p \in P$, denoted by $M[t > M'$. M' is called to be directly reachable from M denoted by $M \rightarrow M'$. Let \rightarrow^* be the reflexive and transitive closure of \rightarrow . A marking M' is reachable from M , iff $M \rightarrow^* M'$. The reachability set of PN is defined by $R(PN) := \{M \mid M_0 \rightarrow^* M\}$. PN is called bounded, iff $\forall p \in P : \exists k \in \mathbb{N}_0 : \forall M \in R(PN) : M(p) \leq k$. PN is called live, iff $\forall t \in T, M \in R(PN) : \exists M' \in R(PN) : M \rightarrow^* M'$ and $M'[t >$. A marking $M \in R(PN)$ is a home state, iff $\forall M' \in R(PN) : M' \rightarrow^* M$. PN is an extended free choice net (EFC Net), iff $\forall (p, t) \in P \times T : I^-(p, t) \leq 1, I^+(p, t) \leq 1$ and $\forall p, p' \in P : p\bullet \cap p'\bullet = \emptyset \vee p\bullet = p'\bullet$.

The notions *deadlock* and *trap* are important for determining the liveness of special classes of Place/Transition Nets, like e.g. EFC Nets. If a deadlock becomes empty, it will remain empty; if a trap becomes marked, it will remain marked. For EFC Nets the liveness property can be characterized using the deadlock and trap notions. Utilizing this characterization it can also be decided whether an EFC net possesses any home states, without generating the reachability set.

Theorem 1 (Liveness and home states in EFC Nets [14, 30]) *Let*

$PN = (P, T, I^-, I^+, M_0)$ be an EFC Net.

- PN is live iff every deadlock contains a marked trap in M_0 .
- If PN is bounded¹ and live the reachability set has home states.

Additional efficient analysis techniques for Place/Transition Nets include the analysis of invariants [22] and certain reduction methods [12]. In order to better cope with graphical complexity, Petri Nets with individual tokens have been introduced. In this paper we will employ Coloured Petri Nets as defined by K. Jensen (cf. [20]):

Definition 3 (CPN) *A Coloured Petri Net (CPN) is a 6-tuple $CPN = (P, T, C, I^-, I^+, M_0)$, where*

- P is a finite and non-empty set of places,
- T is a finite and non-empty set of transitions,
- $P \cap T = \emptyset$,
- C is a colour function defined from $P \cup T$ into non-empty sets,

¹Boundedness is often checked by calculation of S-invariants.

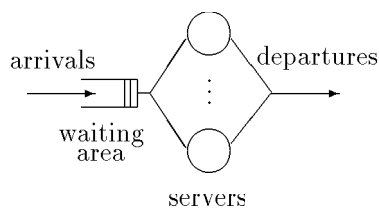


Figure 1: A queue (station)

- I^- and I^+ are the backward and forward incidence functions defined on $P \times T$ such that $I^-(p, t), I^+(p, t) \in [C(t) \rightarrow C(p)_{MS}], \forall (p, t) \in P \times T$.
The subscript MS denotes multi-sets. $C(p)_{MS}$ denotes the set of all finite multi-sets of $C(p)$,
- M_0 is a function defined on P describing the initial marking such that $M_0(p) \in C(p)_{MS}, \forall p \in P$.

As pointed out below, every CPN can be unfolded into an uniquely determined Place/Transition Net. Consequently, all definitions and notions for Place/Transition Nets can also be defined for CPNs, employing this transformation.

Definition 4 (Unfolding of CPNs) *The unfolding of a CPN $= (P, T, C, I^-, I^+, M_0)$ into a PN is performed as follows:*

1. $\forall p \in P, c \in C(p)$ create a place (p, c) of PN.
2. $\forall t \in T, c' \in C(t)$ create a transition (t, c') of PN.
3. Define the incidence functions of PN as

$$I^-((p, c)(t, c')) := I^-(p, t)(c')(c)^2$$

$$I^+((p, c)(t, c')) := I^+(p, t)(c')(c).$$
4. The initial marking of PN is given by $M_0((p, c)) := M_0(p)(c)$.

$PN = (\bigcup_{p \in P} \bigcup_{c \in C(p)} (p, c), \bigcup_{t \in T} \bigcup_{c' \in C(t)} (t, c'), I^-, I^+, M_0)$ is the unfolded CPN.

A well founded theory exists for the qualitative analysis of CPNs: In [19] the reduction methods for Place/Transition Nets [12] are ported to CPNs. Analysis of invariants of CPNs is described in [18], and an efficient technique for constructing the reachability graph is presented in [16], exploiting symmetries in the CPN.

As mentioned in the introduction, pure Petri Nets lack (deliberately!) the notion of time, such that merely certain qualitative features of a system can be modelled and analysed. For performance evaluation purposes other model worlds must obviously be used.

2.2 Queueing Networks

A standard Queueing Network (e.g. [27]) - where we take "standard" to denote the most easily analysable class of networks - comprises a set of interconnected queues. Every queue (or: station, cf. figure 1) represents a service centre where customers may receive

² $I^-(p, t)(c')$ is a multi-set of $C(p)_{MS}$ and $I^-(p, t)(c')(c)$ denotes the number of occurrences of the element c in this multi-set.

service from servers. A station contains one or several servers. A customer arriving at a station will immediately be served if a free server can be allocated to him. Otherwise, the arrived customer will either have to wait for a server allocation, in some waiting area, or he will be served immediately at the expense of preempting one of the customers just receiving service, who will in turn have to wait until later. A station-specific scheduling strategy determines the rules (options, order etc.) of serving customers. Typical strategies include FCFS (first come first served), a rather natural policy, or RR (round robin), a frequently employed time-slicing policy. Work conservative scheduling strategies are assumed throughout, which implies that the server is never idle if customers are in the queue, and that customers do not leave the queue before having been completely served. The service that a customer demands from a station is specified by the amount of time it would take the station to perform this service, given that customer were the only one present. Rather than describing service durations individually, for every customer, service durations of an entire class of customers are specified by a corresponding random variable (more specifically, by its distribution). Different customer classes, as distinguished by different demand distributions, may be acknowledged by a station. Frequently employed demand distributions include the continuous exponential and, more generally, phase-type (Erlang, Coxian) distributions³.

The interconnection of stations is usually but implicitly specified by the routes that customers take among these stations. Again, customer routes are not normally described as pertaining to an individual customer. Rather, the routing behaviours of an entire group of customers, of a so-called customer chain, are specified stochastically by a set of discrete random variables which together prescribe all routing probabilities $r_{i,s;j,t}$, with which customers, having been served at station i according to demand class s , will proceed to station j with demand class t . The notions of classes and chains can be utilised in different ways. Suffice it to say that multiple chains (i.e., several different groups of customer behaviour) can be attained by specifying multiple sets of (station, class)-pairs which are not mutually reachable (i.e., not mutually connected by paths of non-zero routing probabilities). With reference to reflecting an environment of the network or not, a chain can be open or closed. In the open chain case, customers can join the chain from the environment, the arrival instances being specified by a (usually continuous, almost consistently exponential) random variable which describes the duration(s) of time intervals between any two successive arrivals. In this case, possibilities of exiting from the chain, from some station(s) to the environment, ought to reasonably also exist; formal specification may, to cover these aspects, employ a pseudo-station "0" (representing the environment) and enhance the set of routing probabilities according with probabilities $r_{0;j,t}$ for splitting the arrival stream up into (station i , class t)-arrivals and values $r_{i,s;0}$ for denoting the probability of chain exits from (station j , class s)-visits. For the closed chain case, customer arrivals and departures do not exist, resulting in a fixed (and to be specified) number of customers circling in such a chain. A Queueing Network with exclusively open or closed chains is termed an open or closed network, respectively. The coexistence of open and closed chains determines a so-called mixed network. Obviously, many conceivable facets of station and customer behaviours are not covered by these standard Queueing Networks. Those non-covered aspects include typically, but not exclusively, a multitude of non-local aspects such as

³The usual notation for isolated queues is $A/B/m$ -*scheduling strategy*, where A denotes the probability distribution function (pdf) specifying the interarrival times of customers, B is the pdf of service times and m is the number of servers.

state-dependent routing, blocking phenomena, splitting and forking/joining of behaviour processes, synchronisation conditions in general. A variety of non-standard Queueing Networks (Extended QNs) (e.g. [25]) have consequently been devised in order to be at least descriptively capable of coping with one or the other of these aspects.

Analysis of Queueing Networks is typically pursued with the objective of assessing their performance properties. Corresponding measures include station and network populations, station and network throughputs, station utilisations and, most importantly, customer residence times at stations, in subnetworks, in chains (i.e., in the network). The major analysis methodology (as far as practical applicability is concerned) consists of mapping the specified Queueing Network (i.e., the combined stations/chains-complex) onto a corresponding Markov process and subsequently evaluating this process with respect to its transient or, more often, its stationary distribution (provided always that the QN description permits this mapping, that the process allows for stationarity, etc.). In the framework of this methodology, a large subset of the sketched standard QNs, the so-called product form class (e.g. [24]) exhibits analytic solutions for its stationary population distributions which are, moreover, amenable to comparatively extremely efficient evaluation algorithms. Beyond that product form class, and also covering the majority of Extended QNs with finite population state spaces as far as the Markovian conditions are satisfied, numerical evaluation options for Markov chains exist and have in practice been applied for problems of respectable sizes [15]. Yet more general QNs must be attacked by discrete event simulation and corresponding statistical techniques.

2.3 Timed and Stochastic Petri Nets

Approaching the objective of quantitative analysis in the Petri Net framework, several approaches were undertaken in the last two decades. The most important and very popular trend is to directly add time and probability characterizations into Petri Nets [1, 2, 3, 4, 5, 21].

Two principal possibilities exist of integrating timing aspects into Petri Nets:

- specification of a residence time for tokens in a place (Timed places Petri Nets; TPPNs) [26, 31],
- specification of a firing delay for enabled transitions (Timed transition Petri Nets; TTPNs) [1, 21].

In the recent past, the major preference is on TTPNs. TTPNs can be classified into two categories regarding the manner of handling enabled transitions. For TTPNs with *preselection policy*, input tokens needed for firing are destroyed by the transition before the firing phase is started. After a prescribed firing delay, the firing phase terminates by creating the output tokens. The firing process of a transition in TTPNs with preselection, thusly, is not atomic like that of ordinary Petri Nets. If several transitions are in conflict, the one to destroy its tokens first is chosen randomly. Alternatively, for TTPNs with *race policy*, all enabled transitions compete with each other for their input tokens, and the fastest transition will fire first ("race"). Here, firing remains an atomic action.

The most important representatives of TTPNs are those Stochastic Petri Nets, e.g. Generalized Stochastic Petri Nets (GSPNs) [2, 3], that describe Markov processes. Due to the widespread use of GSPNs, we add their particular definition as a representative of Stochastic Petri Nets with race policy. GSPNs consider two types of transitions: timed

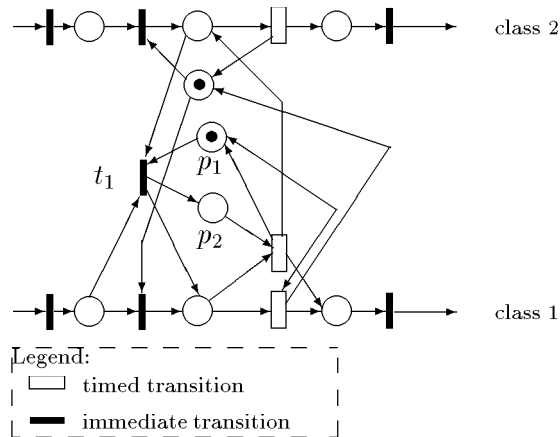


Figure 2: -/M/1-”preemptive-resume priority” station modelled by a GSPN

transitions, firing a specific time after enabling, and immediate transitions, firing immediately (i.e. in zero time).

Definition 5 (GSPN, [1])⁴ A GSPN is a 6-tuple $GSPN=(P,T,I^-,I^+,M_0,W)$ where

- $PN=(P,T,I^-,I^+,M_0)$ is the underlying Place/Transition Net
- $W = (w_1, \dots, w_{|T|})$ ⁵ is an array whose entry w_i
 - is the rate $\in \mathbb{R}^+$ of an exponential distribution as specifying the stochastic firing delay, if transition t_i is a timed transition or
 - is a weight $\in \mathbb{R}^+$ as specifying the relative firing frequency, if transition t_i is an immediate transition.

The firings of any enabled, immediate transitions have priority over that of any simultaneously enabled, timed transitions.

Timed and Stochastic Petri Nets offer the possibilities for both qualitative analysis, using efficient techniques from Petri Net theory, and for quantitative analysis, using numerical Markov techniques or simulation.

Their main disadvantage (as mentioned in the introduction) is their awkwardness in describing scheduling strategies. The following example is meant to illustrate this point. Figure 2 shows a GSPN model of a station serving two classes of customers according to a ”preemptive-resume priority”⁶ - strategy, using standard Petri net elements. Transition t_1 is enabled if preemption is to occur. Firing of t_1 preempts customers of class 2. The markings of p_1 and p_2 keep track of the occurrence of such a situation. This is the reason for modelling the service of a customer of class 1 by two identically parameterized timed transitions. Note that this realization of a preemptive priority scheduling strategy with resumption is only correct if exponentially distributed service times are assumed; otherwise a yet more complex model would have to be devised.

⁴For simplicity we omit the definition of inhibitor arcs and priority levels.

⁵ $|T|$ denotes the number of elements in set T.

⁶A newly arriving customer with higher priority (class 1) preempts the service of the customer being served at that moment. The preempted customer will continue his task, at the point of interruption, after completion of service of all customers with higher priority.

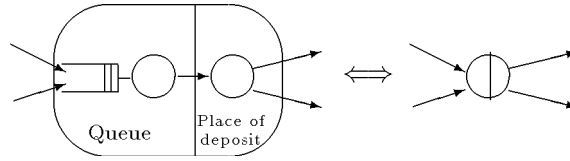


Figure 3: Timed place of a QPN

3 The QPN world

3.1 Definition

In the QPN world [7, 11] timing aspects are added to the places of a (Coloured) Petri Net (cf. figure 3) in a non-standard, structured manner: A queue (station) may be integrated into a place. A corresponding timed place consists of two components, the queue (station) and a "place of deposit" for tokens (customers) having completed their service at this queue⁷. The behaviour of the net is as follows. Tokens, when fired onto a timed place by any of its input transitions, are inserted into the queue according to the station's scheduling strategy. Tokens in a queue are not available for the QPN transitions. After completion of its service, a token (customer) is moved to the "place of deposit". Tokens on this "place" are available for all output transitions of the timed place. An older version of the QPN world [7, 11] forced all transitions to fire immediately after being enabled. In the version described here we also integrate standard timed transitions, obtaining yet more convenient modelling capabilities. An enabled timed transition will fire after a certain delay⁸ specified by a random variable. Enabled immediate transitions will fire according to relative firing frequencies. We assume that tokens are neither generated nor destroyed in queues. As a consequence, qualitative analysis can partially be achieved by analysing the underlying Coloured Petri Net⁹.

Definition 6 (QPN) A *Queueing Petri Net (QPN)* is an 8-tuple $QPN=(P, T, C, I^-, I^+, M_0, Q, W)$ where

- $CPN=(P, T, C, I^-, I^+, M_0)$ is the underlying Coloured Petri Net
- $Q = (q_1, \dots, q_{|P|})$ is an array whose entry q_i
 - denotes the description of a queue taking all colours of $C(p)$ into consideration, if p_i is a timed place or
 - denotes the value (keyword) *untimed*, if p_i is an untimed place.
- $W = (w_1, \dots, w_{|T|})$ is an array of functions whose entry w_i is defined on $C(t_i)$ and $\forall c \in C(t_i) : w_i(c)$ is
 - the description of a distribution specifying the firing delay, if transition t_i is a timed transition w.r.t. colour $c \in C(t_i)$ or
 - a weight $\in \mathbb{R}^+$ specifying the relative firing frequency, if transition t_i is an immediate one w.r.t. colour $c \in C(t_i)$.

⁷This notion of a place should not be confused with the term "place" in Petri Nets.

⁸We assume the race policy here.

⁹This also implies a finite state space of the stochastic process described by the QPN, if the underlying CPN is bounded.

The firings of immediate transitions have priority over those of timed transitions. The graphical representation of transitions and untimed places is similar to that of GSPNs; a pictorial representation of a timed place is given in the right part of figure 3.

Each QPN describes a stochastic process. The state space of this stochastic process is spanned out by the state spaces of all queues of timed places and by the marking spaces of all places of the net. There are two type of states (cf. [2]):

- **tangible states:** If no immediate transition is enabled, only timed actions can occur, either due to the firing of a timed transition or due to the change of a queue state. If all timed actions are specified by exponential (or phase-type) random variables, the stochastic process can be a Markov process.
- **vanishing states:** If an immediate transition is enabled, its firing has priority over all timed actions. The stochastic process leaves such a state immediately; in the case of multiply enabled, immediate transitions, the one to fire is chosen according to the relative firing frequencies.

The initial state of the stochastic process can be constructed using M_0 in the following way. In case of an untimed place, the marking is given as in an ordinary Coloured Petri Net. In case of a timed place, only the "place of deposit" (see figure 3) is marked with the tokens specified by M_0 , all queues are considered empty. Qualitative notions, like boundedness, liveness and the existence of home states (cf. definition 2) can be directly defined for the QPN world in the context of the state space of this stochastic process.

3.2 Examples

The QPN world is a very powerful, albeit easy to handle and understand, model world. Some examples will demonstrate its usefulness.

Figure 4 presents a central server model with memory constraints:

QPN=(P,T,C, I^- , I^+ , M_0 ,Q,W) where

- CPN=(P,T,C, I^- , I^+ , M_0) is the underlying (Coloured) Petri Net¹⁰ given in figure 4,
- Q = (untimed, -/M/ ∞ -IS, -/M/1-PS, untimed, -/M/1-FCFS, -/M/1-FCFS)¹¹,
- $W = (w_1, \dots, w_{|T|})$ is the array of functions whose entry w_i is defined on $C(t_i)$ and (for simplicity we define) $\forall c \in C(t_i) : w_i(c) := 1$. Note that all transitions are immediate.

Figure 5 demonstrates that open Queueing Networks with generally distributed inter-arrival times can be modelled by QPNs.

Figure 6 depicts that time-outs can easily be described with QPNs. After some processing time the host either deposits a message in the buffer (with probability w_2) or continues with a local task (with probability w_1). The sender takes any message from the buffer and transmits it to the receiver. t_3 represents the timer. t_4 and t_5 represent receipt of an incorrect and correct acknowledgement, respectively.

¹⁰In this example, we consider only one class of customers (colour of tokens), so the CPN is simply determined by the given Place/Transition Net.

¹¹IS (infinite server) assumes that every arriving customer will be immediately receive service by an available server; the waiting area of a station with this scheduling strategy is always empty. PS (processor sharing) is a limit case of the RR strategy; all customers are served simultaneously with $\frac{1}{\text{number_of_customers_actually_present}}$ of a server's capacity.

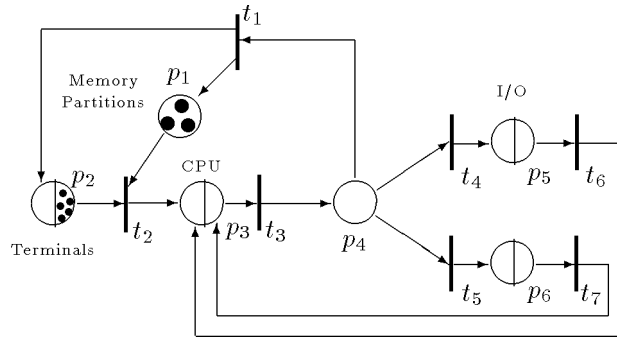


Figure 4: QPN model of a central server model

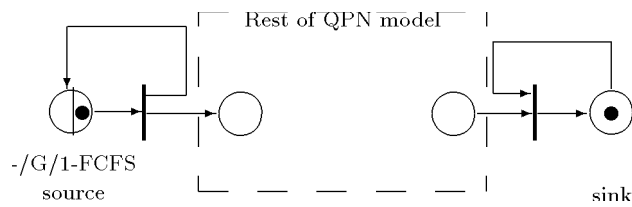


Figure 5: Open Queueing Network modelled by a QPN

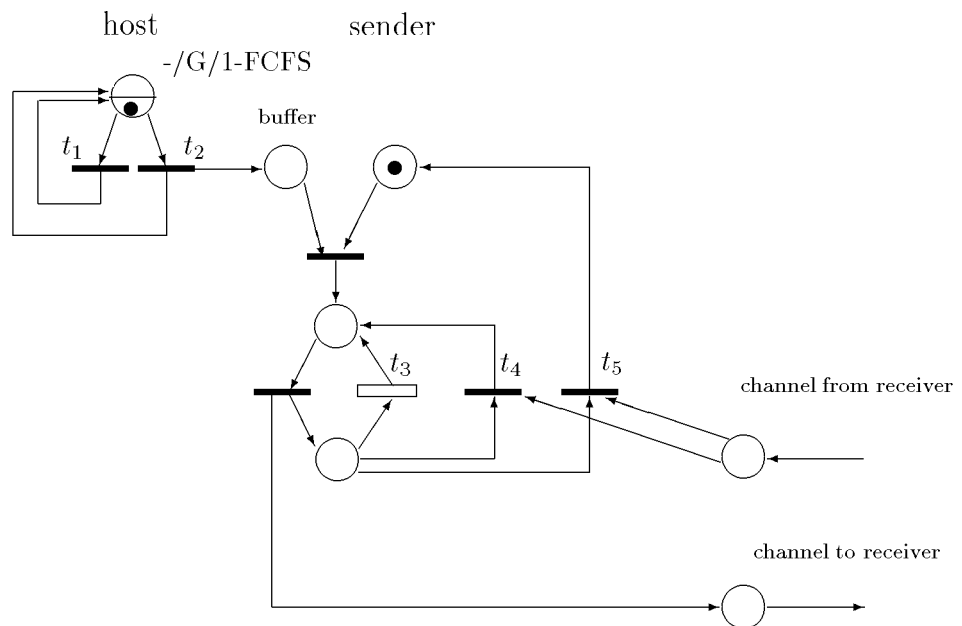


Figure 6: Time-out modelling with QPNs

3.3 Expressive Power

As demonstrated in section 3.2 by way of examples, the QPN world is descriptively more powerful than most of the other model worlds. In [10] it is shown more formally that the QPN world constitutes a superset of (Coloured) Petri Nets, Queueing Networks and several forms of Timed and Stochastic Petri Nets. As an example, we present a construction procedure, proving that every standard Queueing Network can be represented by a QPN:

- For each station i of the Queueing Network create a uniquely determined timed place p_i and an entry q_i with the description of the queue for the array Q .
- For each pair (i,j) of interconnected queues, i.e. \exists classes $s,t: r_{i,s;j,t} > 0$, create a transition $t_{(i,j)}$ of the QPN.

- Define $C: P \cup T \rightarrow \text{set_of_classes}$ by

$$C(x) := \begin{cases} \{t \mid \exists \text{ queue } i, \text{ class } s : r_{i,s;j,t} > 0\} & \text{if } x = p_j \in P \\ \{(s,t) \mid r_{i,s;j,t} > 0\} & \text{if } x = t_{(i,j)} \in T \end{cases}$$

- The incidence functions I^- and I^+ for a pair $(p_i, t_{(j,k)})$ are defined as

$$I^-(p_i, t_{(j,k)})(s,t)(c) := \begin{cases} 1 & \text{if } i = j \text{ and } s = c \\ 0 & \text{otherwise} \end{cases}$$

$$I^+(p_i, t_{(j,k)})(s,t)(c) := \begin{cases} 1 & \text{if } i = k \text{ and } t = c \\ 0 & \text{otherwise} \end{cases}$$

$$\forall c \in C(p_i), (s,t) \in C(t_{(j,k)}).$$

- The initial marking is given by the number of customers in closed chains:

$$M_0(p_i)(c) := \begin{cases} k & \text{if } k \text{ is the number of customers in chain } h \ni c \text{ with} \\ & c \in C(p_i) \text{ and } \forall \text{ queues } j, c' \in C(p_j): \\ & c' \in h, M_0(p_j)(c') \neq 0 \implies (i = j \text{ and } c = c') \\ 0 & \text{otherwise} \end{cases} \quad 12$$

- $Q := (q_1, \dots, q_{|P|})$

- All transitions are immediate, and the array W is given by

$$W := (w_{(1,1)}, w_{(1,2)}, \dots, w_{(1,|P|)}, \dots, w_{(|P|,|P|)}) \text{ where } w_{(i,j)} \text{ is defined as}$$

$$w_{(i,j)}(s,t) := r_{i,s;j,t} \quad \forall (s,t) \in C(t_{(i,j)}).$$

The source and sinks of an open chain can be modelled by the constructs given in figure 5.

4 Combined Qualitative and Quantitative Analysis using QPNs

4.1 A general procedure for the analysis of time-augmented Petri Nets and associated difficulties

As already mentioned, efficient techniques for a qualitative analysis of Petri Nets have been developed in the last decades. Looking at the historical development of time-augmented Petri Nets it seems obvious to attempt a combined qualitative and quantitative analysis using the following procedure (see figure 7): At first, certain qualitative aspects (such as boundedness and liveness) of the Petri Net would be completely analysed, neglecting

¹²Note that this recursive definition does not yield a uniquely determined initial marking for a given Queueing Network; this does, however, not effect the overall behaviour of the net.

the timing aspect. Then, if the Petri net satisfied all qualitative requirements, it would appear worthwhile to perform a quantitative analysis. For this procedure to be justifiable, all qualitative features of the Petri Net would have to remain valid after the introduction of time. Unfortunately, this needs not be the case.

Consider figure 8 where a GSPN is presented with a live underlying Place/Transition Net. Examination of the GSPN's state space yields its non-liveness, since t^* can only fire if places p_1 and p_2 are simultaneously marked. Hence, liveness of a Petri Net can disappear with the incorporation of time.

Similar observations hold for the existence of home states, which is demonstrated for the GSPN world in [8]. In figure 9, the same effect is demonstrated for the QPN world using two classes of customers (colours of tokens). The stations of the timed places p_2 and p_5 serve arriving customers (tokens) employing a FCFS-scheduling strategy. Service times are assumed exponentially distributed. If transition t_1 (or t_3) is enabled, the token of the input place is transferred to place p_2 (or p_5) and a token of the complementary type is put on place p_3 (or p_6). Transitions t_2 and t_4 merge only tokens of the same kind. Scheduling the tokens in a FCFS-manner yields two strongly connected subsets in the state space of the QPN, such that home states do not exist. The subsets are characterized by the sequencing of customers in both queues (either "a" ahead of "b" or "b" ahead of "a"). The main problem for quantitative analyses in such cases is, that the analyst might not be aware of this situation. Using e.g. simulation, only one of these two subsets of the state space is analysed ("by accident"), leading to dangerous misinterpretations of the modelled system (cf. [8]).

Another undesired effect of integrating timing aspects is the occurrence of so-called 'timeless traps' [9]. A typical example of a timeless trap is shown in figure 10. Because of the priority of immediate transitions over all timed actions, transitions t_1 and t_2 will fire forever, with no way out of this situation¹³. The main problem here is that the stochastic process can reach a situation from where all reachable states are vanishing. The description of a corresponding Markov process can hence, not be extracted from the QPN model, although all timing specifications are given by exponential random variables (cf. [9]).

4.2 A solution strategy

All examples of the last section illustrate that the analysis procedure in figure 7 has to be modified or possibly completely changed to yield a useful combination of qualitative and quantitative analysis. What possibilities do we have to cope with this basic problem? One way is to modify existing Petri Net algorithms to render them suitable for analysis of the time-augmented nets; this might turn out a rather uncomfortable approach, as Petri Net theory and existing software products could no longer be used directly. Furthermore, all future research in the Petri Net area would have to be adapted to time-augmented Petri Nets. The other way is to develop suitable restrictions for an integration of timing aspects, such that results of a qualitative analysis remain valid for quantitative analysis. This idea leads to the analysis procedure given in figure 11. A great benefit of proceeding in this manner is that the standard theory of Petri Nets remains unaffected; additionally, any forthcoming results on efficient algorithms for qualitative analyses would be usable

¹³That situation will occur if the consumer has consumed all items of the buffer, and the producer is in his production phase.

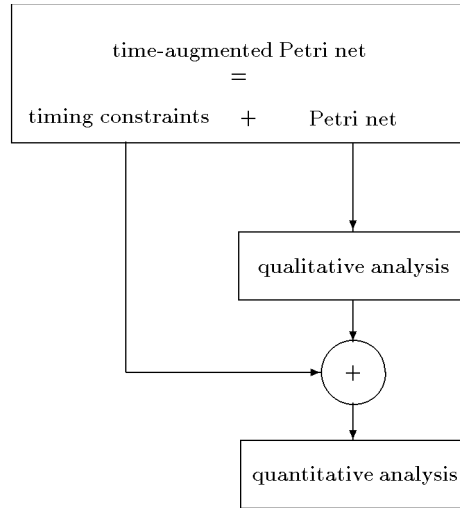


Figure 7: Principal analysis procedure for time-augmented Petri Nets

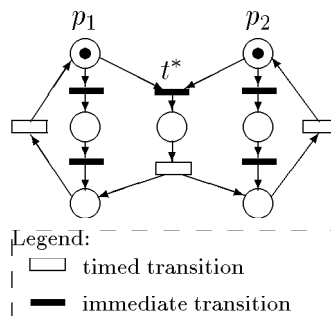
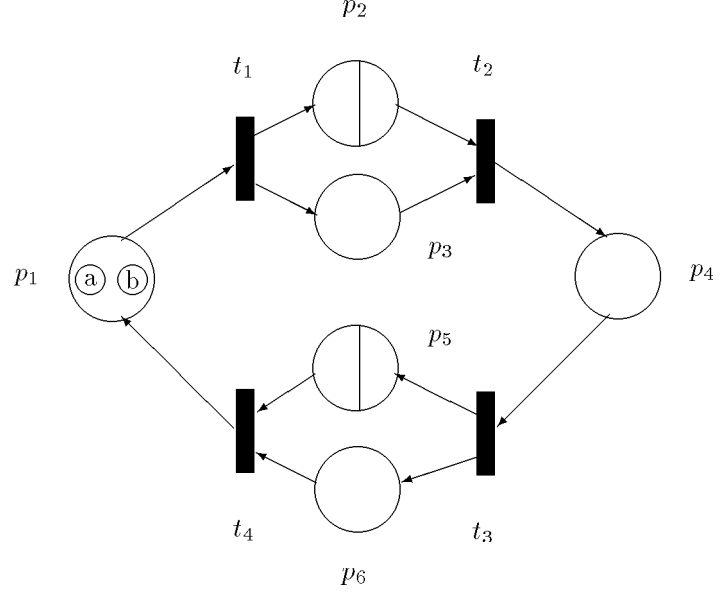


Figure 8: GSPN with dead transitions



$QPN = (P, T, C, I^-, I^+, M_0, Q, W)$ where

- $CPN = (P, T, C, I^-, I^+, M_0)$ is given by

- $P = \{p_1, \dots, p_6\}$

- $T = \{t_1, \dots, t_4\}$

- $C(x) = \{a, b\}, \forall x \in P \cup T$

- $I^-(p_i, t_j)(x)(y) = \begin{cases} 1 & \text{if } (x = y) \text{ and} \\ & [i = 1 \wedge j = 1] \vee [i = 2 \wedge j = 2] \vee \\ & [i = 3 \wedge j = 2] \vee [i = 4 \wedge j = 3] \vee \\ & [i = 5 \wedge j = 4] \vee [i = 6 \wedge j = 4] \\ 0 & \text{otherwise} \end{cases}$

$\forall p_i \in P, t_j \in T, \forall x, y \in \{a, b\}$

- $I^+(p_i, t_j)(x)(y) = \begin{cases} 1 & \text{if } (x = y) \text{ and} \\ & [i = 1 \wedge j = 4] \vee [i = 2 \wedge j = 1] \vee \\ & [i = 4 \wedge j = 2] \vee [i = 5 \wedge j = 3] \\ 1 & \text{if } (x \neq y) \text{ and} \\ & [i = 3 \wedge j = 1] \vee [i = 6 \wedge j = 3] \\ 0 & \text{otherwise} \end{cases}$

$\forall p_i \in P, t_j \in T, \forall x, y \in \{a, b\}$

- $M_0(p_i)(x) = \begin{cases} 1 & \text{if } i = 1 \\ 0 & \text{otherwise} \end{cases} \quad \forall x \in \{a, b\}, p_i \in P$

- $Q = (q_1, \dots, q_6)$ with $q_i = \begin{cases} \text{-/M/1-FCFS} & \text{if } i = 1 \vee i = 5 \\ \text{untimed} & \text{otherwise} \end{cases}$

- $W = (w_1, \dots, w_4)$ with $w_i(x) = 1, \forall i = 1, \dots, 4$ and $x \in \{a, b\}$

Figure 9: QPN with no home states

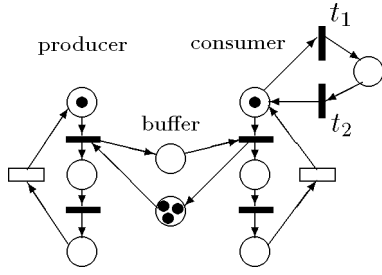


Figure 10: GSPN with a timeless trap

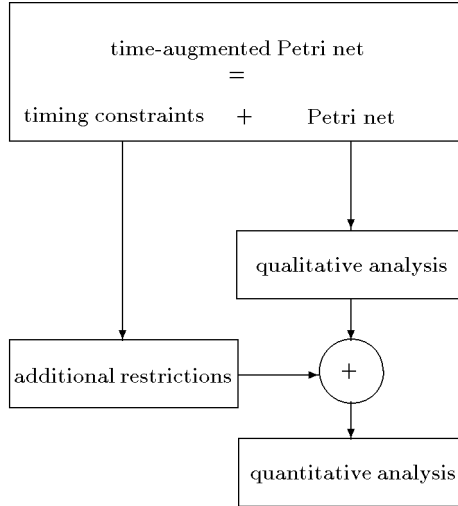


Figure 11: Modified analysis procedure for time-augmented Petri Nets

without difficulty. To find such restrictions for general net structures is difficult. We limit our examination here to EFC Nets. Another reason for this restriction is that this class of Petri Nets has been exhaustively studied in the last decades, and many qualitative features can be characterized by efficiently testable conditions.

In the next section, we will demonstrate that an easily testable restriction guarantees the transfer of all desired features to the QPN, if they are present in the underlying CPN. As the QPN world is a superset of several Stochastic Petri Net worlds [10], these results are valid for these worlds, too.

4.2.1 Extended free choice QPNs

To handle the above mentioned problems, we first consider the single-class case only, i.e. we assume

$$(*) \quad |C(p)| = 1, \forall p \in P.$$

The underlying Petri Net of the QPN can now be defined as a Place/Transition Net in a straightforward way, simplifying our notation. Furthermore we assume:

- that at least one timed transition or one timed place is defined in the QPN¹⁴ and that all timing constraints are specified by continuous probability distributions with unlimited support.

¹⁴otherwise we would only have the description of a Petri Net without timing aspects.

- that all queueing disciplines are work conservative.

Definition 7 (Extended free choice QPN) A QPN is an extended free choice QPN (EFC-QPN), iff the unfolded CPN $PN=(P, T, I^-, I^+, M_0)$ is an EFC net.

Definition 8 (Condition EQUAL-Conflict) A QPN= $(P, T, C, I^-, I^+, M_0, Q, W)$ satisfies condition EQUAL-Conflict iff

$$\begin{aligned} \forall t, t' \in T : \exists p \in P, c \in C(p), c' \in C(t), c'' \in C(t') : \\ I^-(p, t)(c')(c) \times I^-(p, t')(c'')(c) > 0 \implies \\ \text{either } t, t' \text{ are both timed or are both untimed transitions} \\ \text{w.r.t. colours } c' \in C(t), c'' \in C(t'). \end{aligned}$$

Condition EQUAL-Conflict implies that conflicts can only occur between transitions of the same kind.

Assumption (*) ensures that every QPN can be represented by an equivalent single-class QPN using the procedure given in definition 4. If PN is the unfolded CPN, the descriptions of queues are not affected, and the "timing information" for a transition can be directly attached to a transition (t, c') of the Place/Transition Net, if t is a timed or immediate transition w.r.t. colour $c' \in C(t)$. To simplify our notation, we will use this unfolded QPN in the following theorems.

Theorem 2 (Timeless traps in EFC-QPNs) Let QPN= $(P, T, C, I^-, I^+, M_0, Q, W)$ be an EFC-QPN with a bounded and live underlying Petri Net. QPN satisfies condition EQUAL-Conflict \implies QPN does not have any timeless traps.

Proof: Assume a timeless trap exists.

$\implies \exists \tilde{T} \subseteq T : \forall t \in \tilde{T} : t$ is untimed, and $\bullet t$ is a set of untimed places, and the stochastic process can reach a situation where all transitions of \tilde{T} will fire forever, and all transitions of $T \setminus \tilde{T}$ are dead. Let $\tilde{P} := \bullet \tilde{T}$. The assumptions at the beginning of this section ensure $T \setminus \tilde{T} \cup P \setminus \tilde{P} \neq \emptyset$. Boundedness and liveness of the underlying Petri Net yields its strong connectedness ([13]), so that one of the following two cases can occur:

- a) $\exists p \in \tilde{P} : p \bullet \supseteq \{t, t'\}, t \in \tilde{T}, t' \in T \setminus \tilde{T}$.
 t is enabled iff t' is enabled, as the QPN is an EFC Net satisfying condition EQUAL-Conflict. This contradicts our assumption of the existence of a timeless trap (where all transitions of $T \setminus \tilde{T}$ would have to be dead).
- b) $\exists t \in \tilde{T} : t \bullet \supseteq \{p, p'\}, p \in \tilde{P}, p' \in P \setminus \tilde{P}$.
As t can fire forever and $p' \in P \setminus \tilde{P}$, this place is not bounded which contradicts the initial assumption. \square

Condition EQUAL-Conflict is merely sufficient for the avoidance of timeless traps, but not necessary. The next theorem shows that this condition is necessary and sufficient for the liveness of EFC-QPNs.

Theorem 3 (Liveness in EFC-QPNs) Let QPN= $(P, T, C, I^-, I^+, M_0, Q, W)$ be an EFC-QPN with a bounded and live underlying Petri Net. QPN satisfies condition EQUAL-Conflict \iff QPN is live.

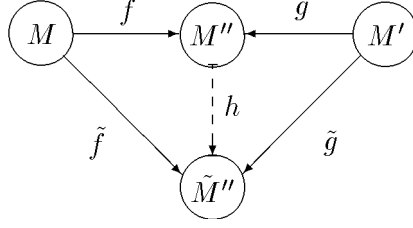


Figure 12: Sketch of proof

Proof:

" \implies :" Assume the QPN is not live. Then the dynamic behaviour of the QPN can reach a situation, where a $t \in T$ exists which will never again be enabled. Therefore, a place $p \in \bullet t$ exists which will never again be marked. This holds due to the EFC net structure and to the work-conservation of all scheduling disciplines. Therefore, all transitions $t' \in \bullet p$ are dead as well.

$\implies \exists S_{empty} \subseteq S, \exists T_{dead} \subseteq T : \bullet S_{empty} \subseteq T_{dead}$ and $S_{empty} \bullet = T_{dead}$ due to the EFC net structure. $\implies \bullet S_{empty} \subseteq S_{empty} \bullet$ and S_{empty} is an unmarked deadlock, which contradicts the liveness of the underlying Petri Net (cf. theorem 1).

" \impliedby :" obvious, as the firings of immediate transitions have priority those of timed ones. \square

A very important result is that condition EQUAL-Conflict also ensures the existence of home states for live and bounded EFC-QPNs. If the QPN is a description of a Markovian process, this implies the existence of a stationary distribution.

Theorem 4 (Home states in EFC-QPNs (cf. [7, 11])) *Let*

QPN = (P, T, C, I⁻, I⁺, M₀, Q, W) be an EFC-QPN with a bounded and live underlying Petri Net.

QPN satisfies condition EQUAL-Conflict \implies QPN has home states.

Sketch of proof:

Because of assumption (*) and due to the work conservation of all scheduling disciplines, it is sufficient to consider a reduced state space description, where a queue state is represented by the number of present customers only. For each state of the QPN's stochastic process, we define a corresponding marking of the underlying CPN by adding the population of a queue and of the corresponding "place of deposit", for all timed places. The set of attained markings is a subset of the reachability set of the underlying CPN. Now assume that the state space of the QPN does not include any home states. This implies that the set of corresponding states has at least two strongly connected disjoint subsets (cf. [30]). Let M and M' be two markings of the two subsets (see figure 12). The reachability set of the EFC Net has home states following theorem 1 (e.g. M'' in figure 12), such that there exist firing sequences $f, g \in T^* : M[f > M''$ and $M'[g > M''$. Let T_{im} be the set of immediate transitions. The QPN's liveness (cf. theorem 3) implies the existence of a firing sequence $h \in T_{im}^*$ such that permutations \tilde{f} and \tilde{g} of fh and gh , respectively, can be constructed with $M[\tilde{f} > \tilde{M}''$ and $M'[\tilde{g} > \tilde{M}''$ and \tilde{f} and \tilde{g} fireable with respect to the QPN's stochastic process. \tilde{M}'' is a member of both subsets contradicting our initial assumption. \square

Up to now we have only considered QPNs with one class of customers (tokens; cf. assumption (*)). The next theorem extends our results provided that all queueing disciplines are of a specific form.

Definition 9 (Condition STATION) *A QPN= $(P, T, C, I^-, I^+, M_0, Q, W)$ satisfies condition STATION iff*

$\forall p_i \in P : q_i$ *is untimed*

or

the scheduling discipline is of the following type:

$\forall c \in C(p) : \text{if customers of this class reside at the station,}$

then at least one of them is served at any point in time.

Disciplines like PS or IS satisfy condition STATION, because all customers are served simultaneously, whereas FCFS does only satisfy this condition if $|C(p)| = 1$. If more than one class of customers exists, customers of a particular class will not be served at a point in time, at which a customer of another class occupies the server.

Theorem 5 (QPNs with several classes of customers (tokens)) *Let*

QPN= $(P, T, C, I^-, I^+, M_0, Q, W)$ be an EFC-QPN with a bounded and live underlying Petri Net.

QPN satisfies condition EQUAL-Conflict and Condition STATION

\implies

- *QPN does not include a timeless trap and*
- *QPN is live and*
- *QPN has home states.*

Proof:

Unfolding the QPN yields a single-class QPN satisfying our assumption (*). As we have not introduced any further assumptions with respect to specific service rates of customers in queues or to firing rates of timed transitions, the theorem follows directly from theorems 2, 3 and 4. \square

5 Conclusions

The QPN world presented in this paper is a superset of Petri Nets, Queueing Networks and several forms of Timed and Stochastic Petri Nets [10]. It offers a convenient and familiar notation for the combined description of qualitative and quantitative aspects of a system. Furthermore, QPNs can serve as a basis for theoretical discussions of this subject.

We have pointed out the basic problems which occur when combining qualitative and quantitative analyses of time-augmented Petri Nets, and we have presented a straightforward method for a useful analysis procedure. For certain classes of QPNs, sufficient and necessary conditions have been given which leave all qualitative features of an untimed model valid for the timed model, too. We believe that the discussion presented in this paper can serve as a starting point for future research, in particular by extending the results for EFC-QPNs to further net classes, e.g. to simple nets.

References

- [1] M. Ajmone-Marsan. *Stochastic Petri Nets: an elementary introduction*. Advances in Petri Nets. G. Rozenberg (ed.), Lecture Notes in Computer Science 424, 1989, pp. 1-29.
- [2] M. Ajmone-Marsan, G. Balbo, G. Conte. *A class of Generalized Stochastic Petri Nets for the performance evaluation of multiprocessor systems*. ACM Transactions on Computer Systems, Vol. 2, No. 2, Mai 1984, pp. 93-122.
- [3] M. Ajmone-Marsan, G. Balbo, G. Conte. *Performance models of multiprocessor systems*. MIT Press, London (Great Britain), 1986.
- [4] M. Ajmone-Marsan, G. Balbo, A. Bobbio, G. Chiola, A. Cumani. *The effect of execution policies in the semantics and analysis of Stochastic Petri Nets*. IEEE Transactions on Software Engineering, Vol. 15, No. 7, 1989, pp. 832-846.
- [5] M. Ajmone-Marsan, G. Balbo, G. Chiola, G. Conte. *Applicability of Stochastic Petri Nets to performance modeling*. Second International Workshop on Applied Mathematics and Performance Reliability. Models of Computer Communication Systems, Italien, may 1987, pp. 517-534.
- [6] G. Balbo, S.C. Bruell, S. Ghanta. *Combining Queueing Networks and Generalized Stochastic Petri Nets for the solution of complex models of system behaviour*. IEEE Transactions on Computers, Vol. 37, No. 10, Oct. 1988, pp. 1251-1268.
- [7] F. Bause. *Funktional und quantitativ analysierbare Rechensystemmodelle; über Verbindungen von Petrinetz- und Warteschlangen-Modellwelten*. Diplomarbeit, Fachbereich Informatik der Universität Dortmund (FRG), 1986.
- [8] F. Bause. *Problems in combining qualitative and quantitative analysis using Generalized Stochastic Petri Nets*. Petri Net Newsletters 36, Gesellschaft der Informatik, Special Interest Group on Petri Nets and related System Models, August 1990, pp. 17-21.
- [9] F. Bause. *No way out ∞ the timeless trap*. Petri Net Newsletters 37, Gesellschaft der Informatik, Special Interest Group on Petri Nets and related System Models, December 1990, pp. 4-8.
- [10] F. Bause. *"QN + PN = QPN". Combining Queueing Networks and Petri Nets*. prepared for publication.
- [11] F. Bause, H. Beilner. *Eine Modellwelt zur Integration von Warteschlangen- und Petri-Netz-Modellen*. Proceedings of the 5. GI/ITG-Fachtagung "Messung, Modellierung und Bewertung von Rechensystemen und Netzen", Braunschweig (FRG), 1989, pp. 190-204.
- [12] G. Berthelot. *Transformations and the decompositions of Nets*. in [23], pp. 359-376.
- [13] E. Best. *Structure theory of Petri Nets: the free choice hiatus*. in [23], pp. 168-205.
- [14] E. Best, J. Desel. *Partial order behaviour and structure of Petri Nets*. GMD, No. 373, Gesellschaft für Mathematik und Datenverarbeitung mbH, Sankt Augustin (FRG), March 1989.
- [15] P. Buchholz. *Die strukturierte Analyse Markovscher Modelle*. Dissertation, Fachbereich Informatik der Universität Dortmund, 1991.
- [16] G. Chiola, C. Dutheillet, G. Franceschinis, S. Haddad. *On well-formed Coloured Nets and their symbolic reachability graph*. Proceedings of the International Conference on Application and Theory of Petri Nets, Paris (France), June 1990, pp. 387-410.
- [17] E. Cinlair. *Introduction to stochastic processes*. Englewood Cliffs, New Jersey (USA), 1975.
- [18] J.M. Couvreur. *The general computations of flows for Coloured Nets*. Proceedings of the International Conference on Application and Theory of Petri Nets, Paris (France), June 1990, pp. 204-223.
- [19] S. Haddad. *Generalization of reduction theory to Coloured Nets*. Proceedings of the 9th European Workshop on Application and Theory of Petri Nets, Venedig (Italy), June 1988.
- [20] K. Jensen. *Coloured Petri Nets*. in [23], pp. 248-299.

- [21] A. Pagnoni. *Stochastic Nets and performance evaluation*. in [23], pp. 460-478.
- [22] J.L. Peterson. *Petri Net theory and the modeling of systems*. Englewood Cliffs, 1981.
- [23] *Petri Nets. Central Models and Their Properties. Advances in Petri Nets*. Proceedings of an Advanced Course, Bad Honnef (FRG), Sept. 1986, Part I, G. Rozenberg (ed.), Lecture Notes in Computer Science 254, 1986.
- [24] C.H. Sauer, K.M. Chandy. *Computer systems performance modeling*. Prentice Hall, New Jersey, 1981.
- [25] C.H. Sauer, E.A. MacNair, J.F. Kurose. *The research queueing package: past, present, and future*. Proceedings of the National Computer Conference, AFIPS Vol. 51, 1982, pp. 273-280.
- [26] J. Sifakis. *Use of Petri Nets for performance evaluation*. Measuring, Modelling and Evaluating Computer Systems, H. Beilner, E. Gelenbe (ed.), 1977, pp. 75-93.
- [27] E. de Souza e Silva, R.R. Muntz. *Queueing Networks: solutions and applications*. Stochastic analysis of computer and communication systems, Hideaki Takagi (ed.), North-Holland, 1990, pp. 319-399.
- [28] M. Silva, R. Valette. *Petri Nets and flexible manufacturing*. Advances in Petri Nets, Lecture Notes in Computer Science 424, 1989, pp. 374-417.
- [29] M. Vernon, J. Zahorjan, E.D. Lazowska. *A comparison of performance Petri Nets and Queueing Network models*. Proceedings on the International Workshop on Modelling Techniques and Performance Evaluation, Paris (France), 1987, pp. 191-202.
- [30] W. Vogler. *Live and bounded free choice Nets have home states*. Petri Net Newsletters 32, Gesellschaft der Informatik, Special Interest Group on Petri Nets and related System Models, April 1989, pp. 18-21.
- [31] C.Y. Wong, T.S. Dillon, K.E. Forward. *Timed places Petri Nets with stochastic representation of place time*. Proceedings of the International Workshop on Timed Petri Nets, Turin (Italy), 1985, pp. 96-103.