(c) Proper but not live.

(d) Live but not proper.

(e) Neither proper nor live.

6. Write a computer program to construct the reachability graph of a bounded, marked Petri net model.
7. Given the reachability graph of a bounded Petri net, write algorithms to determine whether the Petri net is (a) proper and (b) live.
8. Construct a Petri net model of an AMS in which there are two machines M_1 and M_2. Each part has to undergo two operations, one on M_1 and the other on M_2, in any order (M_1 followed by M_2 or vice versa). Assume that a raw part is always available and neglect loading, fixturing, transporting, etc. When a raw part finds both M_1 and M_2 free, it will be assigned to M_1. After finishing the first operation, the part will go to the next machine if it is available; otherwise it waits on the machine (thus blocking the machine) until the other machine becomes free and then moves to the other machine. If a fresh raw part and an in-process part vie for a machine, the latter will be assigned the machine. Construct the reachability graph of the above model and show that it is bounded, proper, and live.
9. Show that the PN model of a transfer line with no buffers in-between is a marked graph.

5.2 STOCHASTIC PETRI NETS

Classical Petri nets are useful in investigating qualitative or logical properties of concurrent systems, such as mutual exclusion, existence and absence of deadlocks, boundedness, and fairness. However, for quantitative performance evaluation, the concept of time needs to be added to the definition of Petri nets. A natural way of introducing time into a PN is based on its interpretation as a system model in which given a state (marking), a certain amount of time must elapse before an event occurs (i.e., a transition fires). The event is the final result of some activity that is performed by the system when it is in the situation specified by the marking. Time is thus naturally associated with transitions, indicating that they can fire some time after they become enabled. The choice of associating time with transitions is the most frequent in the literature on timed PNs. It may, however, be noted that PNs with timed transitions are equivalent to PNs with timed places. Several researchers such as Ramchandani [1973], Sifakis [1977], and Ramamoorthy and Ho [1980] investigated the use of *timed*

Petri nets in which places or transitions were associated with *deterministic* time durations. The analysis of such timed Petri nets (TPNs) is, however, tractable only in the case of special classes such as marked graphs.

The idea of associating *random* time durations was first explored independently by Natkin [1980] and Molloy [1981], and this was the starting point for the emergence of *stochastic Petri nets* (SPNs) and their extensions as a principal performance modeling tool. In this book, the focus will be on SPNs rather than on deterministic timed Petri nets. Also, we shall consider SPNs in which transitions rather than places are associated with times.

Definition: An SPN is a sex-tuple (P, T, IN, OUT, M_0, F) where (P, T, IN, OUT, M_0) is a Petri net and F is a function with domain $(R[M_0]\ X\ T)$, which associates with each transition in each reachable marking, a random variable.

The above is a very general definition of an SPN. We shall call the function F as the *firing function* and the random variable $F(M, t)$ for $M \in R[M_0]$ and $t \in T$ as the *firing time* of transition t in the marking M. Thus the firing time of a transition in an SPN is in general *marking dependent*. In an SPN, when t is enabled in M, the tokens remain in the input places of t during the firing time of t in M. At the end of the firing time, tokens are removed from the input places of t and tokens are deposited in the output places of t. When a transition t gets enabled, we say t *starts firing* and when the firing time has elapsed, we say t *has finished firing* or also often say t *has fired*. It is, however, possible that t gets disabled at some instant of time before finishing firing, due to the firing of a conflicting transition.

In the SPN literature, most often only continuous random variables have been employed. If only continuous random variables are used, a significant ramification will be that no two concurrently enabled transitions of an SPN can finish firing simultaneously. Also, we make the usual simplifying assumption that these random variables are mutually independent.

If, in addition to random firing times, we allow zero firing times, then we have an interesting class of SPNs, which will be discussed in the next section. Also, there have been SPN proposals in which, besides random or zero firing times, deterministic firing times are also allowed.

Definition: Let (P, T, IN, OUT, M_0, F) be an SPN. Let $X(u)$ represent the marking of the SPN at time $u \geq 0$ and let $X(0) = M_0$. Then $\{X(u) : u \geq 0\}$ is a stochastic process which we shall call the *marking process* of the SPN.

Note that the state space of the marking process of an SPN is $R[M_0]$, the reachability set. The basic philosophy underlying the use of various classes of SPNs in performance evaluation is the equivalence of their marking process, under appropriate distributional assumptions, to a Markov or semi-Markov process with discrete state space. The typical steps in SPN-based performance evaluation include: (1) modeling the given system by an SPN, (2) generating the marking process, (3) computing the steady-state probability distribution of the states of the marking process, and (4) obtaining the required performance measures from the steady-state probabilities. All steps in the SPN-based performance evaluation can be automated, and this constitutes an important reason for the popularity of SPN-based performance modeling.

5.2.1 Exponential Timed Petri Nets

The remainder of this section is devoted to the basic class of SPNs in which all firing times are exponentially distributed. For this reason we shall call this class of SPNs exponential timed Petri nets (ETPNs).

Definition: An ETPN is a sex-tuple (P, T, IN, OUT, M_0, F) in which (P, T, IN, OUT, M_0) is a Petri net and the firing function $F : (R[M_0]\ X\ T) \rightarrow \mathbb{R}$ associates to each transition t in each reachable marking M, an exponential random variable with rate $F(M, t)$.

For the sake of convenience, we shall designate each transition in an ETPN as an *exponential transition* and refer to $F(M, t)$ as the *firing rate* of t in M. A transition of an ETPN is usually represented by a rectangular bar in ETPN diagrams.

Example 5.8

Consider the manufacturing system of Example 3.24, which comprises a single NC machine that works on an inexhaustible supply of raw workpieces one at a time and produces finished parts according to *resume policy* or *discard policy*. The sequence of activities in the above system can be modeled by the ETPNs shown in Figure 5.4. The interpretation of the various elements of these ETPN models is provided in Table 5.3. Note that in the initial marking shown, the machine is being set up for the next

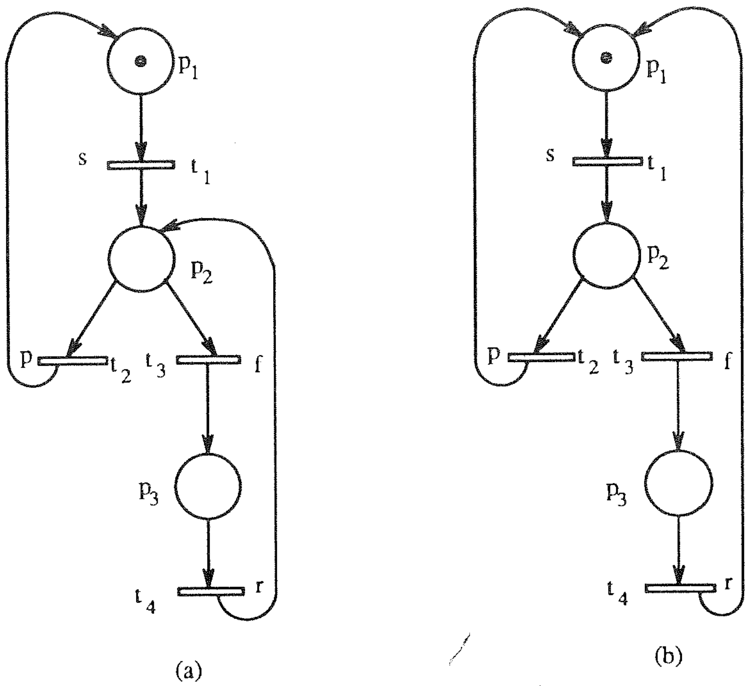


Figure 5.4 ETPN models for resume policy and discard policy

part. For this initial marking, the reachability set of each ETPN comprises three markings, and the reachability graphs are shown in Figure 5.5. The three markings are:

$M_0 = (100)$: Machine being set up for the next part

$M_1 = (010)$: Machine processing a part

$M_2 = (001)$: Machine failed, being repaired.

Note that the above markings are identical to the three states of the Markov chain of Example 3.24.

Equivalence Between ETPNs and CTMCs

Natkin [1980] and Molloy [1981] have shown that the marking process of an ETPN is a continuous time Markov chain (CTMC). This establishes a bridge between SPNs and Markov chains. In the following discussion, we outline a proof for showing the equivalence between an SPN and a CTMC. The proof is presented in the form of two lemmata. In the first lemma,we derive the transition probabilities for the states of the equivalent CTMC.

Table 5.3 Description of the ETPN models of Figure 5.4

Places

1 : Machine being set up for the next operation
2 : Machine processing a workpiece
3 : Failed machine being repaired.

Exponential Transitions

1 : Setup operation
2 : Processing of a part by the machine
3 : Failure of the machine
4 : Repair of the machine.

Firing Rates

$F(M,t_1) = s$; $F(M,t_2) = p$; $F(M,t_3) = f$; $F(M,t_4) = r$

Initial Marking

$p_1 : 1$; $p_2 : 0$; $p_3 : 0$

In the second lemma, we show that the sojourn states in the individual states are exponentially distributed.

Lemma: Let (P, T, IN, OUT, M_0, F) be an ETPN. Given $M_i, M_j \in R[M_0]$, there exists a specific probability a_{ij} of reaching M_j immediately after exiting from M_i.

Proof: Let T_i be the set of enabled transitions in M_i and define:

$$T_{ij} = \left\{ t \in T_i : M_i \xrightarrow{t} M_j \right\}$$

There are two possibilities: $T_{ij} = \phi$ and $T_{ij} \neq \phi$. If T_{ij} is empty, we have that M_j cannot be reached from M_i in a single step and hence $a_{ij} = 0$. Now consider the case when T_{ij} is non-empty. Let

$$\lambda_{ij} = \sum_{t_k \in T_{ij}} F(M_i, t_k), \quad \lambda_i = \sum_{t_k \in T_i} F(M_i, t_k)$$

The probability of marking M_i changing to M_j is the same as the probability that one of the transitions in the set T_{ij} fires before any of

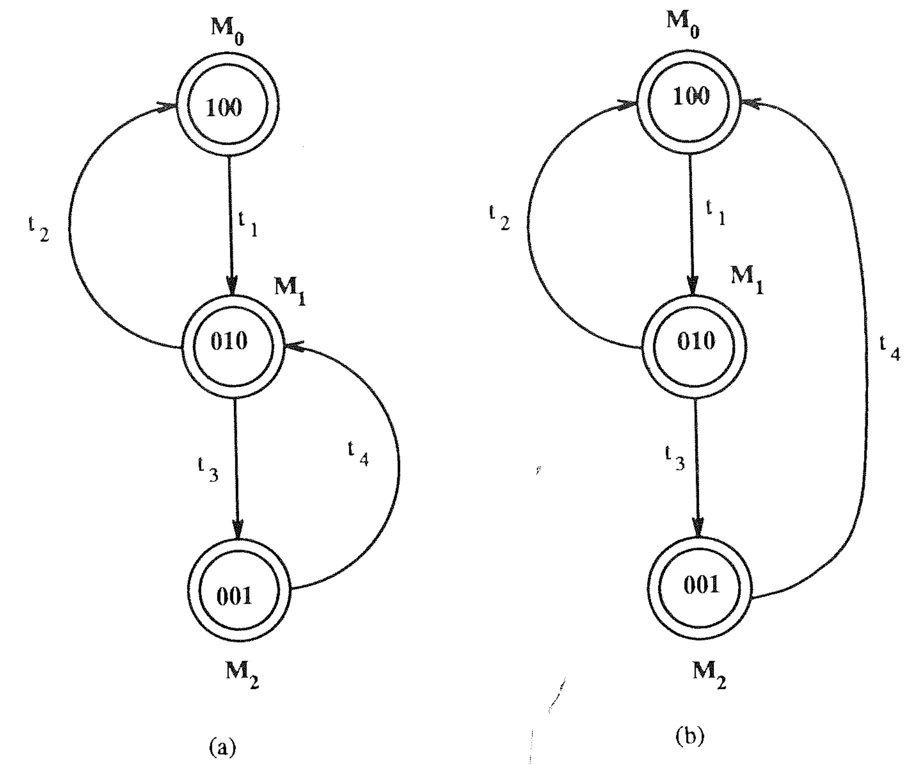


Figure 5.5 Reachability graphs of ETPN models of Figure 5.4

the transitions in the set $T \setminus T_{ij}$. Since the firing times in an ETPN are mutually independent exponential random variables, it follows that the required probability has the specific value given by

$$a_{ij} = \frac{\lambda_{ij}}{\lambda_i}$$

In the expression for a_{ij} deduced above, note that the numerator is the sum of the rates of those enabled transitions in M_i, the firing of any of which changes the marking from M_i to M_j; whereas the denominator is the sum of the rates of all the enabled transitions in M_i. Also note that $a_{ij} = 1$ if and only if $T_{ij} = T_i$.

Example 5.9

Let us look at marking $M_1 = (010)$ in the ETPN model of Figure 5.4(a). Following the notation employed in the lemma above, we have

$$T_1 = \{t_2, t_3\}; \quad T_{10} = \{t_2\}; \quad T_{11} = \phi; \quad T_{12} = \{t_3\}$$

$$\lambda_1 = p + f; \quad \lambda_{10} = p; \quad \lambda_{11} = 0; \quad \lambda_{12} = f$$

Thus we have

$$a_{10} = \frac{p}{p+f}; \quad a_{11} = 0; \quad a_{12} = \frac{f}{p+f}$$

In fact, the matrix of transition probabilities can be seen to be

$$\begin{bmatrix} 0 & 1 & 0 \\ \frac{p}{p+f} & 0 & \frac{f}{p+f} \\ 0 & 1 & 0 \end{bmatrix}$$

Similarly, the matrix of transition probabilities for the ETPN model of Figure 5.4(b) can be verified to be

$$\begin{bmatrix} 0 & 1 & 0 \\ \frac{p}{p+f} & 0 & \frac{f}{p+f} \\ 1 & 0 & 0 \end{bmatrix}$$

Note that the first of the above matrices was derived in Example 3.33.

Lemma: The sojourn time of any reachable marking in an ETPN (P, T, IN, OUT, M_0, F) is exponentially distributed.

Proof: Let $M_i \in R[M_0]$ and let T_i be the set of enabled transitions in M_i. Suppose T_i' is the subset of M_i comprising all transitions, the firing of any of which in M_i would lead to a marking other than M_i. Denote

$$\lambda_i' = \sum_{t_k \in T_i'} F(M_i, t_k)$$

The sojourn time in M_i is a random variable given by

$$\min_{t_k \in T_i'} (EXP(F(M_i, t_k)))$$

Then by the mutual independence of the firing times, it follows that the sojourn time of M_i is exponentially distributed with rate λ_i'.

Example 5.10

Considering again the ETPN models depicted in Figure 5.4, the sojourn times of the markings M_0, M_1, and M_2 can be seen to be exponential random variables with rates $s, p+f$, and r, respectively.

Theorem: The marking process of an exponential timed Petri net is a continuous time Markov chain.

Proof: Proof is immediate on applying the two lemmata above and using the fact that the firing time random variables in an ETPN are all mutually independent.

Note that the state space of the equivalent CTMC is the reachability set $R[M_0]$ of the ETPN. The transition rate from one marking to another can be computed as follows. Let $M_i, M_j \in R[M_0]$. For $M_j \neq M_i$, the transition rate is given by

$$q_{ij} = \sum_{t_k \in T_{ij}} F(M_i, t_k)$$

where T_{ij}, as usual, is the subset of enabled transitions in M_i such that the firing of any transition in T_{ij} leaves the marking process in M_j. For $M_j = M_i$, the transition rate is given by

$$q_{ii} = -\sum_{j \neq i} q_{ij}$$

Example 5.11

The transition rate matrix (infinitesimal generator) of the equivalent CTMC of the ETPNs of Figure 5.4 can be respectively seen to be

$$\begin{bmatrix} -s & s & 0 \\ p & -(p+f) & f \\ 0 & r & -r \end{bmatrix}; \quad \begin{bmatrix} -s & s & 0 \\ p & -(p+f) & f \\ r & 0 & -r \end{bmatrix}$$

Note that the above matrices are identical to the ones in Example 3.28. The performance analysis of these CTMC models has already been discussed in that example.

Example 5.12

Instead of a single machine, we now consider the activities of two identical machines working on an inexhaustible source of raw workpieces. We assume that the setup times, processing times, failure rates, and repair rates for the two machines are the same. To capture the interactions in

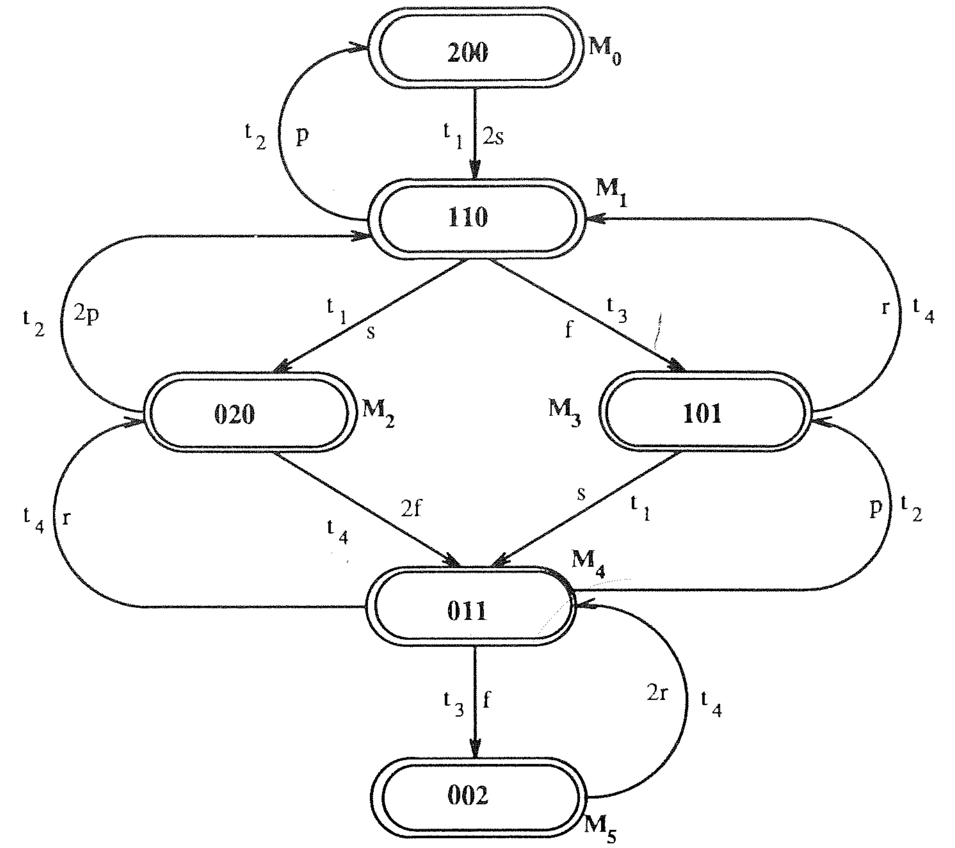


Figure 5.6 Reachability graph with initial marking (200)

this system, we can use the same ETPN models (Figure 5.4) with two changes: (i) The initial marking is changed to state (200), meaning that both machines are being set up for their next operations. (ii) The firing rates are marking dependent, and are given for each $M \in R[M_0]$ by

$$F(M, t_1) = sM(p_1)$$
$$F(M, t_2) = pM(p_2)$$
$$F(M, t_3) = fM(p_2)$$
$$F(M, t_4) = rM(p_3)$$

In the case of t_4, we have assumed that two separate facilities are available to repair both the machines concurrently, if both are in failed condition.

There are six markings in the reachability set now. The transition rate diagram of the equivalent CTMC is shown in Figure 5.6.

PROBLEMS

1. Investigate the complications that will arise in the analysis of SPNs in the following cases:

   (a) Firing times are non-exponential, and

   (b) deterministic and exponential firing times coexist.

2. In discrete time stochastic Petri nets proposed by Molloy [1985], each transition has a firing time that is a geometric random variable. Outline a proof for showing the equivalence between discrete time SPNs and discrete time Markov chains.

3. Investigate whether or not the class of all ETPNs is the same as the class of continuous time Markov chains.

4. Consider a manufacturing system comprising three failure-prone machines and exactly two repair facilities, each of which can attend to one failed machine at a time. Assuming the rest of the details as in Example 5.8, complete the definition of an ETPN for this system. Obtain the reachability graph of the ETPN and the transition probabilities. Also compute the transition rate matrix of the equivalent Markov chain.

5. (Dining Philosophers' Problem). Figure 5.P.2 shows an ETPN model for the classical dining philosophers' problem with three philosophers and three forks. Taking the philosophers as assembly stations and the forks as assem-

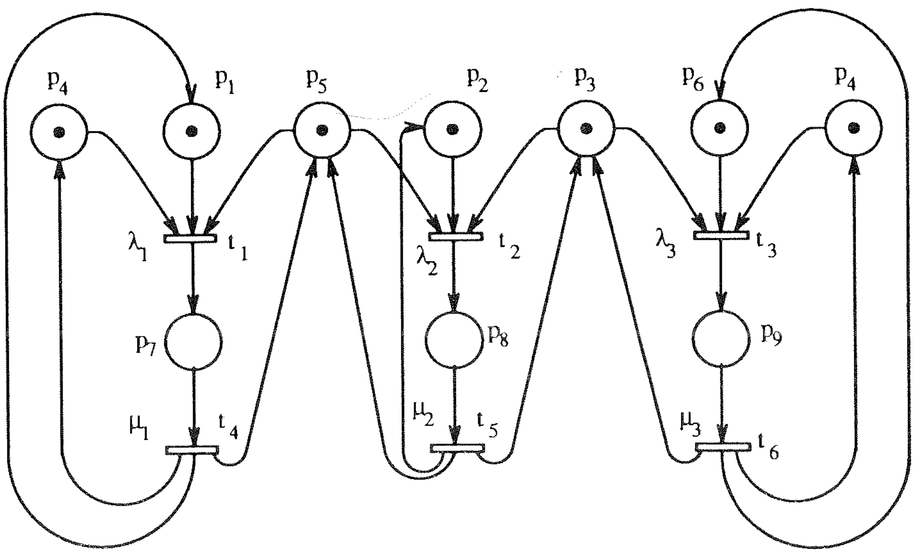


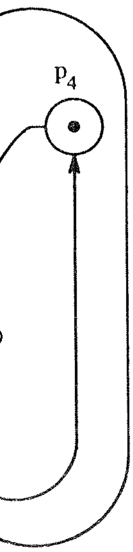
Figure 5.P.2

bling robots, the above problem represents a familiar situation in automated manufacturing. Obtain the reachability graph, transition probabilities, and the infinitesimal generator of the equivalent CTMC.

6. Write a computer program that

   (a) generates the reachability graph of a bounded ETPN, and

   (b) computes the transition rate matrix of the underlying Markov chain.



5.3 GENERALIZED STOCHASTIC PETRI NETS

Generalized stochastic Petri nets (GSPNs), proposed by Ajmone Marsan, Balbo, and Conte [1984], constitute the most extensively used class of SPNs. GSPNs are obtained by allowing transitions to belong to two different classes: *immediate* transitions and *exponential* transitions. Immediate transitions fire in zero time once they are enabled. Exponential transitions have the same significance as in ETPNs. The basic motivation for proposing GSPNs was to avoid having to associate a time with each transition when it is sufficient to associate times only with the activities that are believed to have the largest impact on system performance. A typical example is that of a physical system in which the durations of activities differ by orders of magnitude. In such a case, it is advantageous to model the short activities only from the logical point of view. A strong feature of AMSs that supports the above is the orders of magnitude difference exhibited by the mean times between machine failures, mean processing times, and material handling times. GSPNs are particularly useful because the firing rules are so defined as to reduce the number of states of the associated Markov chain model, thus reducing the solution complexity. Moreover, the availability of a logical structure that can be used in conjunction with the timed one allows the construction of compact performance models of complex systems.

5.3.1 Definition and Firing Rules

In this section we define a GSPN and discuss the transition firing rules to generate the reachability graph.

Definition: A *generalized stochastic Petri net* is an eight-tuple $(P, T, IN, OUT, INH, M_0, F, S)$ where

1. $(P, T, IN, OUT, INH, M_0)$ is an inhibitor-marked Petri net,

2. T is partitioned into two sets: T_I of *immediate* transitions and T_E of *exponential* transitions,
3. $F : (R[M_0] X\ T_E) \rightarrow \mathbb{R}$ is a firing function that associates to each $t \in T_E$ in each $M \in R[M_o]$, an exponential random variable with rate $F(M, t)$,
4. each $t \in T_I$ has zero firing time in all reachable markings, and
5. S is a set, possibly empty, of elements called *random switches*, which associate probability distributions to subsets of conflicting immediate transitions.

In the graphical representation of GSPNs, a horizontal or vertical line represents an immediate transition and a rectangular bar represents an exponential transition. Other conventions regarding graphical representation remain the same.

In a GSPN marking, several transitions may be enabled simultaneously. Let M_i be a reachable marking and let T_i be the set of enabled transitions in M_i. If T_i comprises only exponential transitions, then transition $t_j \in T_i$ fires with probability

$$\frac{F(M_i, t_j)}{\sum_{t_k \in T_i} F(M_i, t_k)}$$

The above situation is identical to the one in any ETPN. If T_i comprises exactly one immediate transition t_j, then t_j is the one that fires. If T_i comprises two or more immediate transitions, then a probability mass function is specified on the set of enabled immediate transitions by an element of S. The firing transition is selected according to this probability distribution. In this case, the set of all enabled immediate transitions, together with the associated probability distribution (switching distribution), is called a *random switch.* The set S is a collection of all random switches of the GSPN model. The probabilities in a switching distribution may be either independent of the current marking, in which case we have a *static* random switch, or dependent on the current marking, in which case we have a *dynamic* random switch.

GSPN markings in which only exponential transitions are enabled are designated as *tangible* markings while the rest of the markings are called *vanishing* markings. Tangible markings represent states in which the system stays for nonzero time whereas vanishing markings are those in which logical changes occur in negligible time.

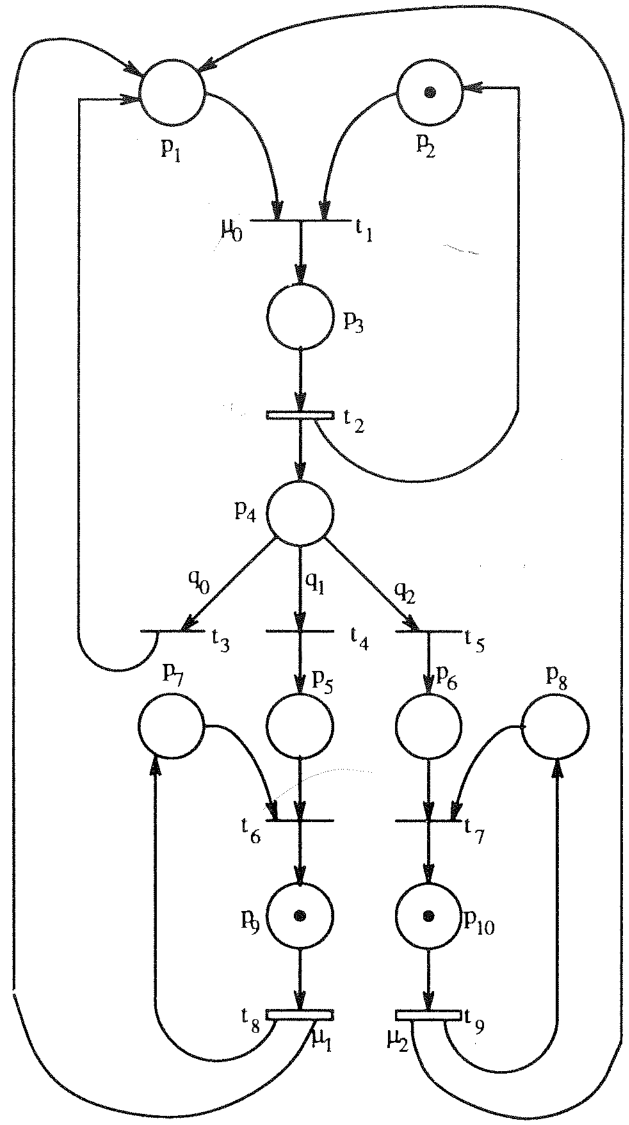


Figure 5.7 GSPN model of a central server FMS

Example 5.13

Consider the closed central server model of a simple FMS that comprises an AGV and two machines M_1 and M_2 (Example 3.25, Figure 3.13). Figure 5.7 depicts a GSPN model of this system assuming that there are two fixtures in the system and that in the initial state, AGV is idle, M_1 is processing one of the parts, and M_2 is processing the other part. Table

Table 5.4 Description of the GSPN model of Figure 5.7

Places

1 : Queue of parts waiting for the AGV
2 : AGV available
3 : AGV transporting a part
4 : A part that has just been transported by the AGV
5 : Queue of parts waiting for M_1
6 : Queue of parts waiting for M_2
7 : M_1 available
8 : M_2 available
9 : M_1 processing a part
10 : M_2 processing a part

Immediate Transitions

1 : AGV starts transporting a part
3 : Finished part gets unloaded from the system
4 : Part joins the queue for M_1
5 : Part joins the queue for M_2
6 : M_1 starts processing a part
7 : M_2 starts processing a part

Random Switch

(t_3, t_4, t_5) with associated probabilities (q_0, q_1, q_2).

Exponential Transitions

2 : Part transfer by the AGV; firing rate = μ_0
8 : Processing by M_1; firing rate = μ_1
9 : Processing by M_2; firing rate = μ_2

Initial Marking

$p_2 : 1$; $p_9 : 1$; $p_{10} : 1$

5.4 gives the interpretation of the elements of this GSPN model. The GSPN model has 10 places, 6 immediate transitions $\{t_1, t_3, t_4, t_5, t_6, t_7\}$, 3 exponential transitions $\{t_2, t_8, t_9\}$ with firing rates μ_0, μ_1, μ_2, respectively, and no inhibitor arcs. The initial marking is $M_1 = (0100000011)$, which we shall denote as $p_2 p_9 p_{10}$ specifying only those places having a token. There is exactly one random switch (static, in this case) comprising the

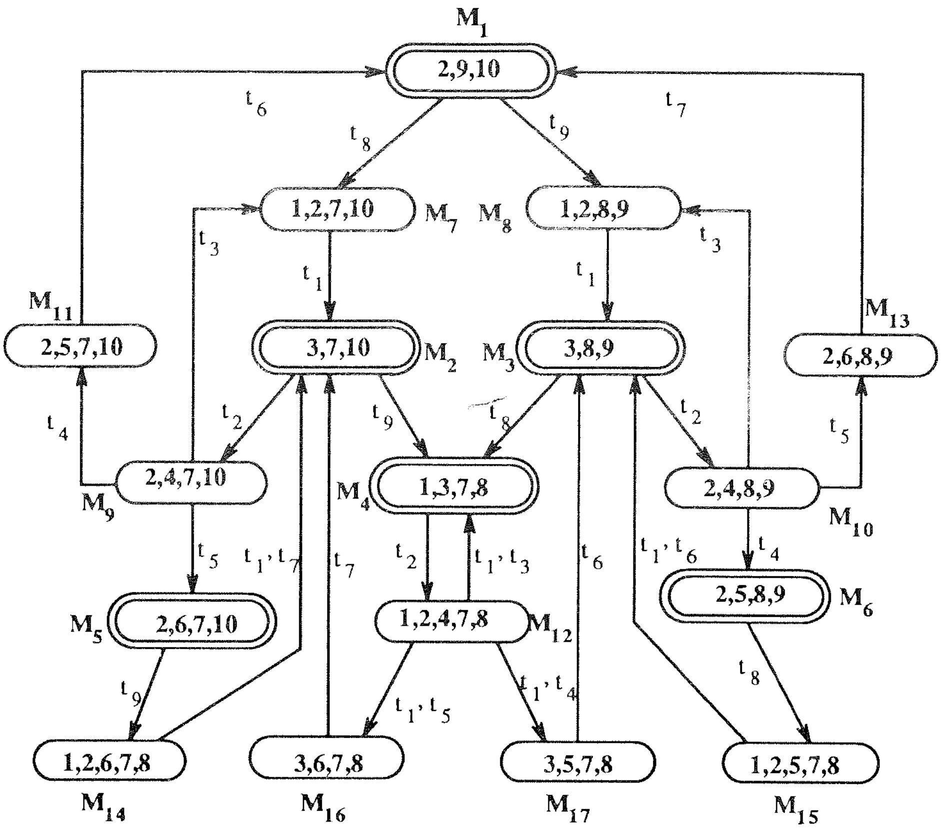


Figure 5.8 Embedded Markov chain for the GSPN model of Figure 5.7

transitions t_3, t_4, and t_5 with corresponding probabilities q_0, q_1, and q_2.

In the initial marking $p_2 p_9 p_{10}$, the exponential transitions t_8 and t_9 are enabled and hence this is a tangible marking. Here t_8 fires with probability $\frac{\mu_1}{\mu_1+\mu_2}$ and t_9 fires with probability $\frac{\mu_2}{\mu_1+\mu_2}$. When t_8 fires, the new marking is $p_1 p_2 p_7 p_{10}$, which is a vanishing marking, as an immediate transition t_1 is enabled in it. Note that t_9 is also enabled here. Only t_1 fires since t_9 is exponential, resulting in the marking $p_3 p_7 p_{10}$, which is a tangible marking. Now t_2 and t_9 are enabled and they can fire with probabilities $\frac{\mu_0}{\mu_0+\mu_2}$ and $\frac{\mu_2}{\mu_0+\mu_2}$, respectively. If t_2 fires, the new marking is $p_2 p_4 p_7 p_{10}$, a vanishing marking enabling t_3, t_4, and t_5. At this stage, the random switch can be invoked to choose the next transition to fire. The evolution of the marking process proceeds as described above and one can construct the reachability graph of the GSPN model in this way. Figure 5.8 shows the reachability graph, with the associated probabilities for transition firings. Vanishing

Table 5.5 Interpretation of the tangible markings in Figure 5.8

M_1	: $p_2p_9p_{10}$: Machines M_1 and M_2 busy; AGV idle
M_2	: $p_3p_7p_{10}$: AGV and M_2 busy; M_1 idle
M_3	: $p_3p_8p_9$: AGV and M_1 busy; M_2 idle
M_4	: $p_1p_3p_7p_8$: AGV busy and the other part waiting for the AGV
M_5	: $p_2p_6p_7p_{10}$: M_2 busy and the other part waiting for M_2
M_6	: $p_2p_5p_8p_9$: M_1 busy and the other part waiting for M_1

markings are shown as single ovals and tangible markings as double ovals. There are 6 tangible markings $\{M_1, ..., M_6\}$ and 11 vanishing markings $\{M_7, ..., M_{17}\}$. The interpretation of the tangible markings is given in Table 5.5. Note that these are identical to the six states given in Example 3.25.

In Figure 5.8, it is also to be noted that in vanishing markings such as M_{12} and M_{14}, two immediate transitions have been simultaneously fired. For example, consider $M_{12} = p_1p_2p_4p_7p_8$. In this marking, t_1, t_3, t_4, and t_5 are the enabled transitions. Among t_3, t_4, and t_5, only one can fire, as determined by the random switch. Transition t_1 is concurrent with each of t_3, t_4, and t_5 and its firing does not in any way affect the latter transitions. Thus we fire the transitions $\{t_1, t_3\}$ with probability q_0, transitions $\{t_1, t_4\}$ with probability q_1, and the transitions $\{t_1, t_5\}$ with probability q_2. Concurrent firing of immediate transitions as detailed above is not always possible and must be decided with care.

5.3.2 Analysis of GSPNs

The marking process of a GSPN $(P, T, IN, OUT, INH, M_0, F, S)$ can be shown to be a semi-Markov process with a discrete state space, given by the reachability set $R[M_0]$. The embedded Markov chain (EMC) of this marking process comprises tangible markings as well as vanishing markings. The transition probability matrix (TPM) of this EMC can be computed using the firing rates and the random switches.

Example 5.14

The reachability graph in Figure 5.8 is essentially the state transition diagram of the EMC of the marking process of the GSPN model being discussed. The transition probabilities for each marking can be computed by examining the set of transitions enabled in the marking. For example, in marking M_1, exponential transitions t_8 and t_9 are enabled and the probability of transiting from M_1 to M_7 is $\frac{\mu_1}{\mu_1+\mu_2}$ and the probability of transiting from M_1 to M_8 is $\frac{\mu_2}{\mu_1+\mu_2}$. In marking M_{11}, only one transition t_6 is enabled and therefore the transition probability from M_{11} to M_1 is 1. As another example, in marking M_9, three (conflicting) immediate transitions t_3, t_4, and t_5 are enabled and by invoking the random switch, the corresponding transition probabilities are fixed as q_0, q_1, and q_2, respectively. It will be instructive to compute and verify all other transition probabilities.

The marking process of a GSPN leaves each vanishing marking as soon as it enters the marking, because an immediate transition fires. Thus sojourn time in each vanishing marking is zero, and consequently from the performance evaluation point of view, it suffices to study the evolution of tangible markings alone. In order to remove vanishing markings from the EMC, a reduced embedded Markov chain (REMC) is defined, including only tangible markings and the transition probabilities in the REMC are deduced from those in EMC as follows:

Let the tangible markings be $M_1, M_2, ..., M_s$ and the vanishing markings be $M_{s+1}, M_{s+2}, ..., M_{s+v}$, where s is the number of tangible markings and v is the number of vanishing markings in the reachability set. The TPM of the EMC can be partitioned as

$$\begin{bmatrix} TT & TV \\ VT & VV \end{bmatrix}$$

where TT gives the one-step transition probabilities from tangible markings to tangible markings, TV gives the one-step transition probabilities from tangible markings to vanishing markings, and so on. If A is the TPM of the REMC, then it can be shown that

$$A = TT + \sum_{k=0}^{\infty} TV * VV^k * VT \qquad (2)$$

where + and * above denote matrix addition and matrix multiplication, respectively. A computationally efficient method for obtaining A has been outlined by Ajmone Marsan, Balbo, and Conte [1984].

Example 5.15

For the GSPN model of Figure 5.7, the REMC is shown in Figure 5.9. The transition probability matrix of the REMC can be computed using

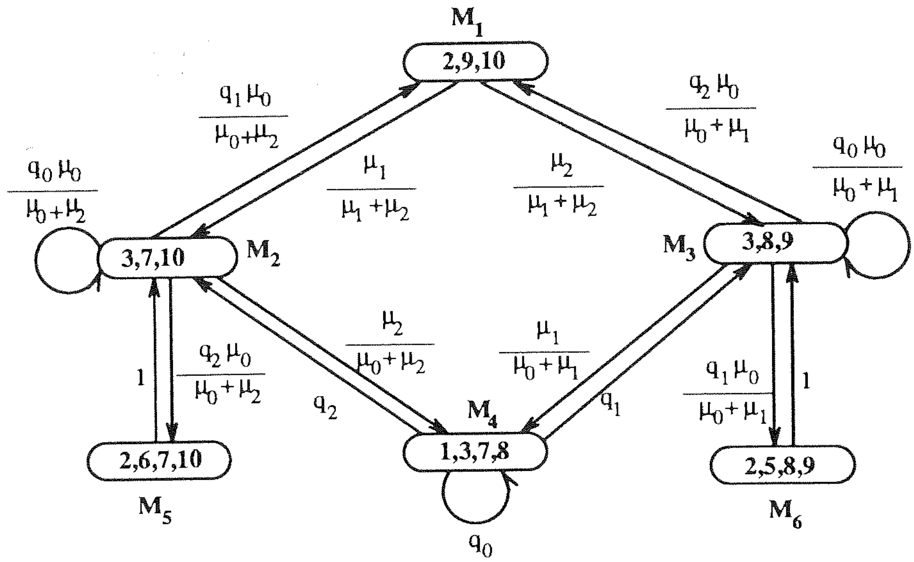


Figure 5.9 Reduced embedded Markov chain for the GSPN model of Figure 5.7

(2). The individual transition probabilities are labeled on the directed arcs connecting the tangible markings.

In order to ensure the existence of a unique steady-state probability distribution for the marking process of GSPNs, the following simplifying assumptions are made:

1. The GSPN is bounded. That is, the reachability set is finite.
2. Firing rates do not depend on time parameters. This ensures that the equivalent Markov chain is homogeneous.
3. The GSPN model is proper and deadlock-free. That is, the initial marking is reachable with a nonzero probability from any marking in the reachability set and also there is no "absorbing" marking.

Most of the GSPN models discussed in this book satisfy the above conditions. However, in Section 5.5 we discuss GSPN models having deadlocks. The above assumptions make the marking process of a GSPN, a finite state space, homogeneous, irreducible, and positive recurrent Markov process. The steady-state probability distribution can therefore be computed as follows:

Let $Y = (y_1, y_2, ..., y_s)$ be a vector of real numbers. Then the solution of the following equations

$$YA = Y$$
$$\text{and} \sum_{i=1}^{s} y_i = 1 \tag{3}$$

gives the stationary probabilities of the REMC. y_i gives the relative number of visits to M_i by the marking process. To obtain the steady-state probabilities of the marking process, we then use the expression

$$\pi_i = \frac{y_i m_i}{\sum_{j=1}^{s} y_j m_j}; \qquad i = 1, 2, ..., s \tag{4}$$

where π_i is the steady-state probability of marking M_i in the marking process (proportion of time the marking process spends in M_i) and m_i is the mean sojourn time of the marking M_i. It has been shown in Section 5.2 that

$$m_i = \frac{1}{\sum_{t_k \in T_i} F(M_i, t_k)} \tag{5}$$

where T_i is the set of enabled transitions in M_i.

Example 5.16

It is easy to verify that the GSPN model of Figure 5.7 satisfies assumptions 1, 2, and 3 above. Hence, unique steady-state probabilities exist. The mean sojourn times are given, using (5), by

$$m_1 = \frac{1}{\mu_1 + \mu_2}; \quad m_2 = \frac{1}{\mu_0 + \mu_2}; \quad m_3 = \frac{1}{\mu_0 + \mu_1};$$
$$m_4 = \frac{1}{\mu_0}; \quad m_5 = \frac{1}{\mu_2}; \quad m_6 = \frac{1}{\mu_1}$$

We can now get the steady-state probabilities by first solving for $y_1, y_2, ..., y_6$ using (3) and then substituting in (4). These probabilities are next used in computing various performance measures. It may be noted that the steady-state probability of each vanishing marking is identically equal to zero.

5.3.3 Computation of Performance Measures

Let $(P, T, IN, OUT, INH, M_0, F, S)$ be a GSPN with the set of reachable tangible markings given by $\{M_1, M_2, ..., M_s\}$, with steady-state probabilities $(\pi_1, \pi_2, ..., \pi_s)$. Certain generic mean performance measures can be defined as shown below. The discussion holds for an ETPN model also and in general for any SPN model.

1. Probability that a particular condition C holds:

$$PROB(C) = \sum_{j \in S_1} \pi_j \tag{6}$$

where $S_1 = \{j \in \{1, 2, ..., s\} : C$ is satisfied in marking $M_j\}$

2. Probability that a place p_i has exactly k tokens $(k = 0, 1, 2, ...)$:

$$PROB(p_i, k) = \sum_{j \in S_2} \pi_j \tag{7}$$

where $S_2 = \{j \in \{1, 2, ..., s\} : M_j(p_i) = k\}$

3. Expected number of tokens in a place p_i:

$$ET(p_i) = \sum_{k=1}^{K} kPROB(p_i, k) \tag{8}$$

where K is the maximum number of tokens p_i may contain in any reachable marking.

4. Throughput rate of an exponential transition t_j :

$$TR(t_j) = \sum_{i \in S_3} \pi_i F(M_i, t_j) q_{ij} \tag{9}$$

where $S_3 = \{i \in \{1, 2, ..., s\} : t_j$ is enabled in $M_i\}$, and q_{ij} is computed as follows. $q_{ij} = 1$ if t_j is not in conflict with any of the enabled transitions in M_i. Otherwise, q_{ij} is the probability that t_j fires among the conflicting enabled transitions in M_i.

5. Throughput rate of immediate transitions:
The throughput rates of immediate transitions can be computed from those of the exponential transitions and the structure of the GSPN model. This will be illustrated in Example 5.17.

6. Mean waiting time in a place p_i:
This can be computed by invoking Little's result as

$$WAIT(p_i) = \frac{ET(p_i)}{\sum_{t_j \in IT(p_i)} TR(t_j)} = \frac{ET(p_i)}{\sum_{t_j \in OT(p_i)} TR(t_j)} \qquad (10)$$

Recall in the above that $IT(p_i)$ is the set of input transitions of p_i and $OT(p_i)$, the set of output transitions of p_i.

Example 5.17

We illustrate the computation of some typical performance measures for the GSPN model of Figure 5.7.

1. Probability that AGV queue is empty

$$= PROB(p_1, 0) = \pi_1 + \pi_2 + \pi_3 + \pi_5 + \pi_6$$

2. Mean length of AGV queue = $PROB(p_1, 1) = \pi_4$
3. Mean number of customers in the AGV queue and the AGV (or the AGV subsystem) = $(\pi_2 + \pi_3) + 2\pi_4$
4. Utilization of AGV = $PROB(p_3, 1) = PROB(p_2, 0) = \pi_2 + \pi_3 + \pi_4$
5. Throughput rates of exponential transitions
$TR(t_2) = \pi_2\mu_0 + \pi_3\mu_0 + \pi_4\mu_0 = (\pi_2 + \pi_3 + \pi_4)\mu_0$
$TR(t_8) = \pi_1\mu_1 + \pi_3\mu_1 + \pi_6\mu_1 = (\pi_1 + \pi_3 + \pi_6)\mu_1$
$TR(t_9) = (\pi_1 + \pi_2 + \pi_5)\mu_2$
6. The throughput rates of the immediate transitions can be computed as:

$$TR(t_1) = TR(t_2); \quad TR(t_3) = q_0\, TR(t_2)$$
$$TR(t_4) = q_1\, TR(t_2); \quad TR(t_5) = q_2\, TR(t_2)$$
$$TR(t_6) = TR(t_8); \quad TR(t_7) = TR(t_9)$$

7. Mean waiting time in AGV queue:

$$WAIT(p_1) = \frac{ET(p_1)}{TR(t_3) + TR(t_8) + TR(t_9)}$$

8. Production rate of parts is given by $TR(t_3)$, the throughput rate of t_3.
9. Mean MLT is given by $\frac{2}{TR(t_3)}$, since the population of the network is 2.

5.3.4 Representational Power of GSPN Models

In this section we show the flexibility and power of GSPN models by presenting three examples. The models presented show how priorities, dynamic routing schemes, and changes in the architecture of the physical system can be captured by GSPNs. The discussion is again centered on the central server FMS with one AGV and two machines.

Example 5.18

In the FMS being discussed, let us make the following change in decision rules: If M_1 and M_2 are both available, the waiting part is assigned to M_2 (probably because M_2 is faster). If M_1 alone is available or M_2 alone is available, the part is assigned to the available machine. Finally, if both M_1 and M_2 are currently busy, the part is assigned to one of the queues probabilistically, as in the previous case. The existence of priorities and conditional decisions make the CQN model non-product form and it may be required to use approximate techniques to find the solution. GSPNs can, however, be used to model the new decision rule exactly. Figure 5.10 shows the new GSPN model. Note that this model has two additional places p_{11} and p_{12} and five additional transitions $t_{10}, ..., t_{14}$, compared to the original model (Figure 5.7). The interpretation of these is included in Table 5.6. The transition t_5 of the previous model, however, is missing in the new model in order to facilitate the modeling of the new decision rules. The priorities manifest in these decision rules are captured by the three inhibitor arcs from p_7 to t_{11}, p_8 to t_{11}, and p_8 to t_{10}. For example, t_{10} fires only if the waiting part finds a token in p_7 (that is, M_1 is available) and there is no token in p_8 (that is, M_2 is not available). Further, the lone random switch in the previous model is now replaced by two random switches in the new model, since the probabilistic routing to M_1 or M_2 arises only in the event that both the machines are busy. Table 5.6 also shows the probability assignments to these random switches.

Figure 5.11 shows the reachability graph of the new GSPN model with the same initial marking $M_1 = (0100000011)$ as before. It is interesting to note that the number of tangible markings is now 4 (compared to 6 in the previous model) and the number of vanishing markings is 9 (compared to 11 previously). The reduction in the number of markings is obviously due to the new decision rules. The analysis of this GSPN model and performance evaluation can be carried out in the same way as detailed before. It would be instructive to compute the TPM of the EMC, the TPM of the REMC, the

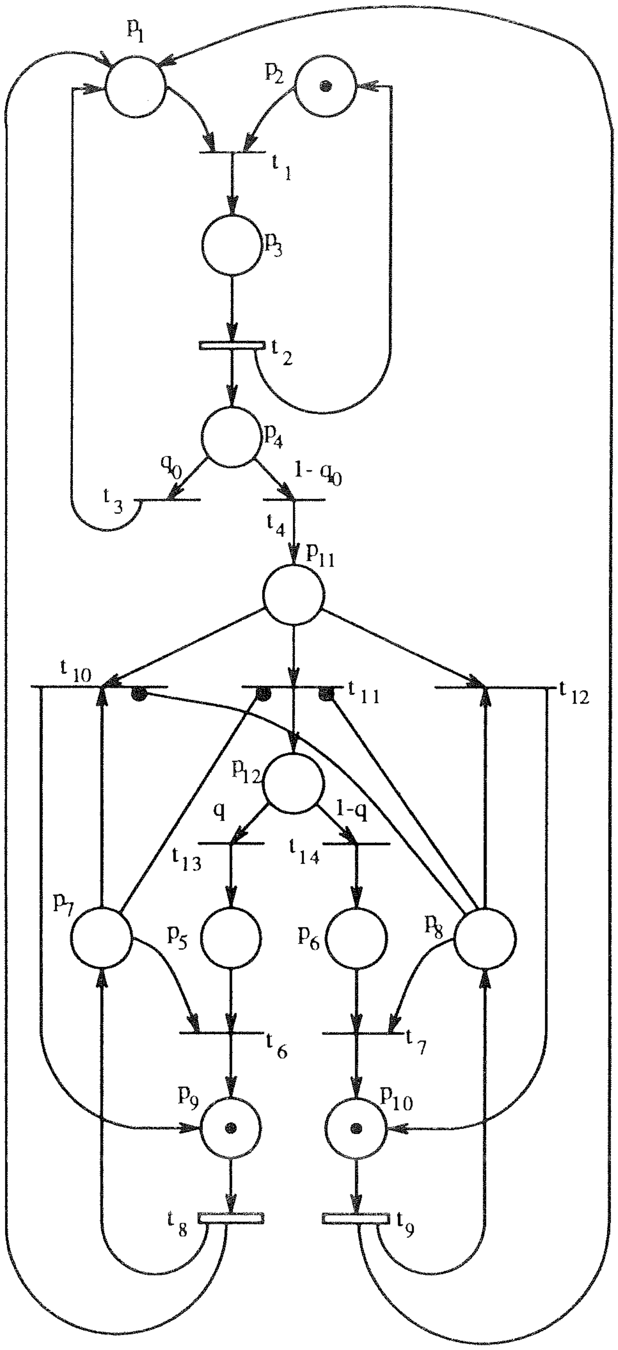


Figure 5.10 A GSPN that models priorities

Table 5.6 Legend for Figure 5.10

Places

11 : A part that will be scheduled onto M_1 or M_2
12 : A part that finds both the machines M_1 and M_2 busy

Transitions

10 : M_1 available while M_2 is busy
11 : M_1 and M_2 both busy
12 : M_2 available
13 : A part is assigned to the M_1 queue
14 : A part is assigned to the M_2 queue.

Inhibitor Arcs

p_7 to t_{11}; p_8 to t_{11}; p_8 to t_{10}

Random Switches

$[t_3 : \; q_0; \quad t_4 : \; 1 - q_0]$

$[t_{13} : \; q; \quad t_{14} : \; 1 - q]$ where $q = \dfrac{q_1}{q_1 + q_2}$ and $q_0 + q_1 + q_2 = 1$

steady-state probabilities, and different performance measures to compare the performance of the two systems discussed so far.

Example 5.19

Here we modify further the decision rule of Example 5.18. If a part finds both machines busy, it will be assigned to the queue that has fewer waiting parts. Further, if the two queues comprise the same number of waiting parts, the routing will be based on the probabilities q_1 and q_2. The above decision rule can be captured by the GSPN model of Figure 5.10, by making the random switch $[t_{13}, t_{14}]$ dynamic. Consider the following definitions for the switching probabilities:

$$\begin{aligned} Prob\,(t_{13}) &= 0 && \text{if } M(p_5) > M(p_6) \\ &= \frac{q_1}{q_1 + q_2} && \text{if } M(p_5) = M(p_6) \\ &= 1 && \text{if } M(p_5) < M(p_6) \end{aligned}$$

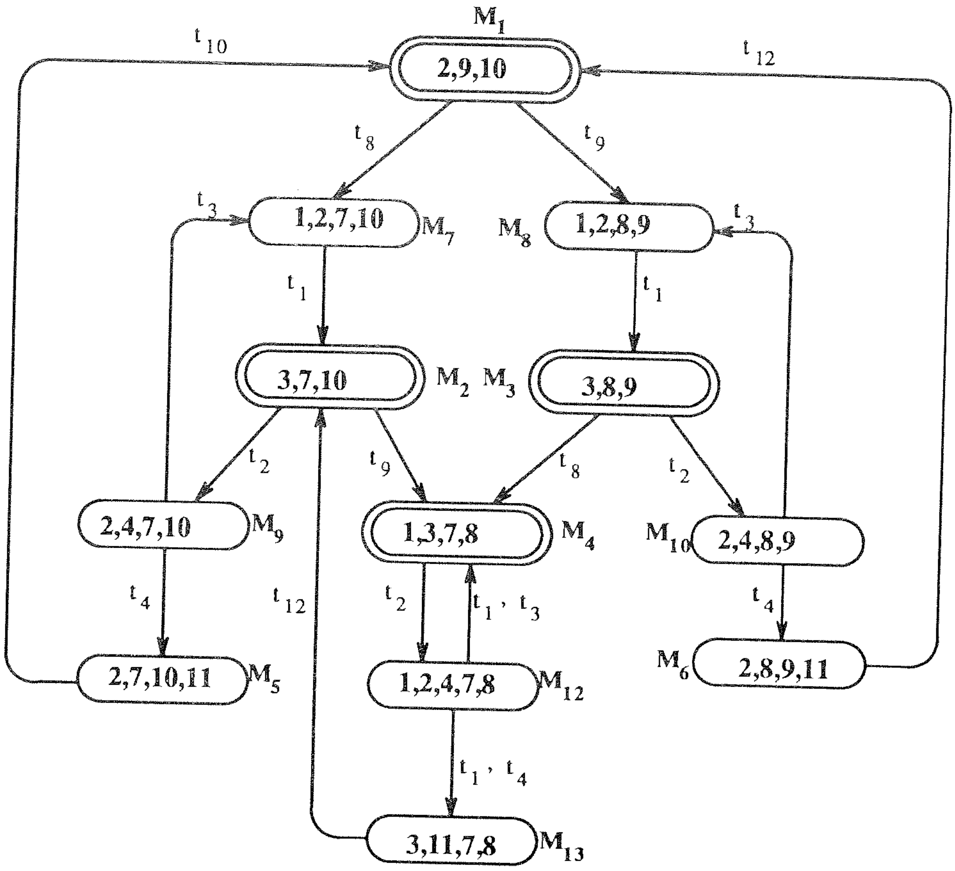


Figure 5.11 Embedded Markov chain of the modified GSPN model

$$
\begin{aligned}
Prob\,(t_{14}) &= 0 && \text{if } M\,(p_6) > M\,(p_5) \\
&= \frac{q_2}{q_1 + q_2} && \text{if } M\,(p_6) = M\,(p_5) \\
&= 1 && \text{if } M\,(p_6) < M\,(p_5)
\end{aligned}
$$

In the above definitions, M is the *current* marking of the GSPN. The probabilities are now being assigned to t_{13} and t_{14} in accordance with the latest decision rule. The above example shows the use and power of dynamic random switches as a modeling paradigm in GSPNs.

Example 5.20

Let us say the number of fixtures in the system is increased to seven. With respect to the GSPN models of Figures 5.7 and 5.10, consider the following initial marking:

$$p_1 : 2; \quad p_3 : 1; \quad p_5 : 1; \quad p_6 : 1; \quad p_9 : 1; \quad p_{10} : 1$$

The above refers to the situation where the AGV, M_1, and M_2 are all busy; two parts are waiting in the AGV queue, and one part each is waiting in the M_1 and M_2 queues. This is one manifestation of the FMS system with seven fixtures. Because of the irreducibility of the underlying Markov chain, there are several other initial markings that can represent the given system.

In addition, let us say we have two identical M_1 machines and two identical M_2 machines instead of one of each type in the previous case. Then the initial marking

$$p_1 : 1; \quad p_3 : 1; \quad p_5 : 1; \quad p_9 : 2; \quad p_{10} : 2$$

will yield the desired GSPN model. Of course, here again there are several other initial markings that can generate the same reachability graph, because of the reversibility of this GSPN model. There is another important change that needs to be done in this case, namely in the firing rates of t_8 and t_9. To account for the simultaneous activity of the two machines of each type, we define the following *dynamic* (or *marking-dependent*) firing rates:

$$t_8 : \; M(p_9)\mu_1 \quad \text{and} \quad t_9 : \; M(p_{10})\mu_2$$

Note that the changes effected as shown in this example significantly vary the reachability set, EMC, REMC, and the steady-state probabilities, but the new values can be computed in an automated way.

PROBLEMS

1. Figure 5.P.3 shows a part of a GSPN model comprising immediate transitions t_1, t_2, t_3. Can t_1, t_2 be fired simultaneously? Obtain sufficient conditions under which two or more immediate transitions can be concurrently fired in a GSPN.
2. Show that the steady-state probability distribution of the tangible markings of the GSPN is the same as that of a CTMC embedded within the marking process. From the EMC and the REMC, how can one obtain the transition rate matrix of the CTMC?
3. Derive the expression

$$A = TT + \sum_{k=0}^{\infty} TV * VV^k * VT$$