Stochastic Petri Net

Serge Haddad

LSV
ENS Cachan, Université Paris-Saclay & CNRS & INRIA
haddad@lsv.ens-cachan.fr

Petri Nets 2016, June 20th 2016

Plan

Stochastic Petri Net versus Time Petri Net

- In TPN, the delays are *non deterministically* chosen.
- In Stochastic Petri Net (SPN), the delays are *randomly* chosen by sampling distributions associated with transitions.

... but these distributions are not sufficient to eliminate non determinism.

Policies for a net

One needs to define:

- The *choice* policy.
  What is the next transition to fire?
- The *service* policy.
  What is the influence of the enabling degree of a transition on the process?
- The *memory* policy.
  What become the samplings of distributions that have not be used?

Stochastic Petri Net versus Time Petri Net

- In TPN, the delays are *non deterministically* chosen.
- In Stochastic Petri Net (SPN), the delays are *randomly* chosen by sampling distributions associated with transitions.

... but these distributions are not sufficient to eliminate non determinism.

Policies for a net

One needs to define:

- The *choice* policy.
  What is the next transition to fire?
- The *service* policy.
  What is the influence of the enabling degree of a transition on the process?
- The *memory* policy.
  What become the samplings of distributions that have not be used?

Choice Policy

In the net, associate a distribution D_i and a weight w_i with every transition t_i.

Preselection w.r.t. a marking m and enabled transitions T_m

- Normalize weights w_i of the enabled transitions: $w'_i \equiv \frac{w_i}{\sum_{t_j \in T_m} w_j}$
- Sample the distribution defined by the w'_i's.
- Let t_i be the selected transition, sample D_i giving the value d_i.

versus

Race policy with postselection w.r.t. a marking m

- For every $t_i \in T_m$, sample D_i giving the value d_i.
- Let T' be the subset of T_m with the smallest delays. Normalize weights w_i of transitions of T': $w'_i \equiv \frac{w_i}{\sum_{t_j \in T'} w_j}$
- Sample the distribution defined by the w'_i's yielding some t_i.

Priorities between transitions could added to refine the selection.

Choice Policy

In the net, associate a distribution D_i and a weight w_i with every transition t_i.

Preselection w.r.t. a marking m and enabled transitions T_m

- Normalize weights w_i of the enabled transitions: $w'_i \equiv \frac{w_i}{\sum_{t_j \in T_m} w_j}$
- Sample the distribution defined by the w'_i's.
- Let t_i be the selected transition, sample D_i giving the value d_i.

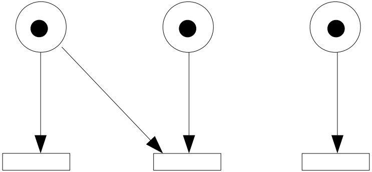
versus

Race policy with postselection w.r.t. a marking m

- For every $t_i \in T_m$, sample D_i giving the value d_i.
- Let T' be the subset of T_m with the smallest delays.
  Normalize weights w_i of transitions of T': $w'_i \equiv \frac{w_i}{\sum_{t_j \in T'} w_j}$
- Sample the distribution defined by the w'_i's yielding some t_i.

Priorities between transitions could added to refine the selection.

Choice Policy: Illustration



$t_1(D_1, w_1)$ $t_2(D_2, w_2)$ $t_3(D_3, w_3)$

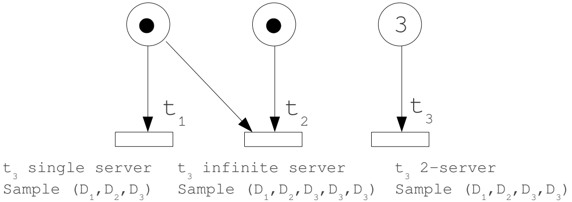
$w_1 = 1$ $w_2 = 2$ $w_3 = 2$

Preselection	Race Policy
Sample (1/5,2/5,2/5)	Sample (D_1, D_2, D_3)
Outcome $\mathbf{t_1}$	Outcome (**3.2**,6.5,**3.2**)
Sample D_1	Sample(1/3,-,2/3)
Outcome **4.2**	Outcome $\mathbf{t_3}$

Server Policy

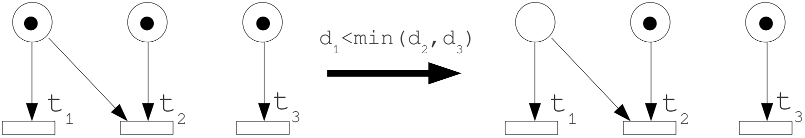
A transition t can be viewed as server for firings:

- A *single server* t allows a single instance of firings in m if $m[t\rangle$.
- An *infinite server* t allows d instances of firings in m where $d = \min(\left\lfloor \frac{m(p)}{Pre(p,t)} \right\rfloor \mid p \in {}^{\bullet}t)$ is *the enabling degree*.
- A *multiple server* t with bound b allows $\min(b, d)$ instances of firings in m.



This can be generalised by marking-dependent services.

Memory Policy (1)



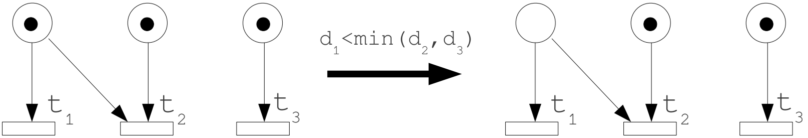
What happens to d_2 and d_3?

Resampling Memory

Every sampling not used is forgotten.

This could correspond to a “crash” transition.

Memory Policy (2)



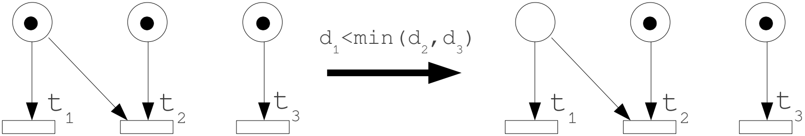
What happens to d_2 and d_3?

Enabling Memory

- The samplings associated with still enabled transitions are kept and decremented ($d_3' = d_3 - d_1$).
- The samplings associated with disabled transitions are forgotten (like d_2).

Disabling a transition could correspond to abort a service.

Memory Policy (3)



What happens to d_2 and d_3?

Age Memory

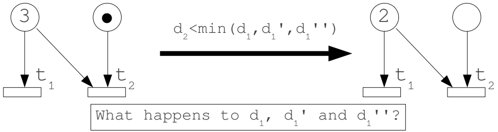
- All the samplings are kept and decremented ($d_3' = d_3 - d_1$ and $d_2' = d_2 - d_1$).
- The sampling associated with a disabled transition is frozen until the transition become again enabled (like d_2').

Disabling a transition could correspond to suspend a service.

Memory Policy (4)

Specification of memory policy

To be fully expressive, it should be defined w.r.t. any pair of transitions.



Interaction between memory policy and service policy

Assume enabling memory for t_1 when firing t_2 and infinite server policy for t_1.
Which sample should be forgotten?

- The last sample performed,
- The first sample performed,
- The greatest sample, etc.

Warning: This choice may have a critical impact on the complexity of analysis.

Plan

Stochastic Petri Net

2 Markov Chain

Markovian Stochastic Petri Net

Generalized Markovian Stochastic Petri Net (GSPN)

Product-form Petri Nets

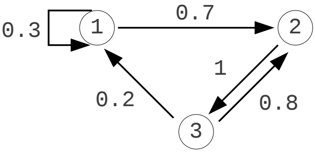
Discrete Time Markov Chain (DTMC)

A DTMC is a stochastic process which fulfills:

- For all n, T_n is the constant 1
- The process is *memoryless*

$$\begin{aligned} Pr(S_{n+1} = s_j \mid S_0 = s_{i_0}, ..., S_{n-1} = s_{i_{n-1}}, S_n = s_i) \\ = Pr(S_{n+1} = s_j \mid S_n = s_i) \\ \equiv \mathtt{P}[i, j] \end{aligned}$$

A DTMC is defined by S_0 and $\mathtt{P}$



$$\mathbf{P} \begin{pmatrix} 0.3 & 0.7 & 0.0 \\ 0.0 & 0.0 & 1.0 \\ 0.2 & 0.8 & 0.0 \end{pmatrix}$$

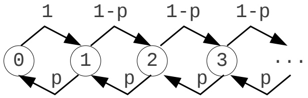
Analysis: the State Status

The *transient analysis* is easy and effective in the finite case:
$\pi_n = \pi_0 \cdot \mathrm{P}^n$ with π_n the distribution of S_n

The *steady-state analysis* ($\exists? \lim_{n\to\infty} \pi_n$) requires theoretical developments.

Classification of states w.r.t. the asymptotic behaviour of the DTMC

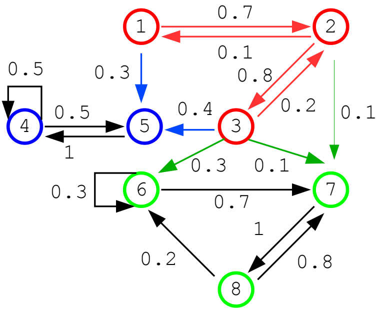
- A state is *transient* if the probability of a return after a visit is less than one. Hence the probability of its occurrence will go to zero. $(p < 1/2)$
- A state is *recurrent null* if the probability of a return after a visit is one but the mean time of this return is infinite.
  Hence the probability of its occurrence will go to zero. $(p = 1/2)$
- A state is *recurrent non null* if the probability of a return after a visit is one and the mean time of this return is finite. $(p > 1/2)$



State Status in Finite DTMC

In a finite DTMC

- The status of a state only depends on the graph associated with the chain.
- A state is transient iff it belongs to a non terminal *strongly connected component* (scc) of the graph.
- A state is recurrent non null iff it belongs to a terminal scc.



$T = \{1, 2, 3\}$

$C_1 = \{4, 5\}$

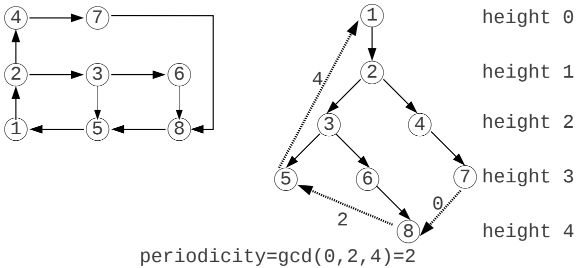
$C_2 = \{6, 7, 8\}$

Analysis: Irreducibility and Periodicity

Irreducibility and Periodicity

- A chain is *irreducible* if its graph is strongly connected.
- The *periodicity* of an irreducible chain is the greatest integer p such that: the set of states can be partionned in p subsets $\mathcal{S}_0, \dots, \mathcal{S}_{p-1}$ where every transition goes from $\mathcal{S}_i$ to $\mathcal{S}_{i+1\%p}$ for some i.

Computation of the periodicity

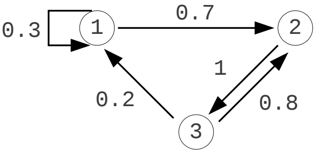


Analysis of a DTMC: a Particular Case

A particular case

The chain is irreducible and *aperiodic* (i.e. its periodicity is 1)

- $\pi_\infty \equiv \lim_{n\to\infty} \pi_n$ exists and its value is independent from π_0.
- π_∞ is the unique solution of $X = X \cdot \mathtt{P} \wedge X \cdot \mathbf{1} = 1$
  where one can omit an arbitrary equation of the first system.



$$\boldsymbol{\pi}_\infty = \begin{pmatrix} 1/8 & 7/16 & 7/16 \end{pmatrix}$$

$$\pi_1 = 0.3\pi_1 + 0.2\pi_2 \quad \pi_2 = 0.7\pi_1 + 0.8\pi_3 \quad \pi_3 = \pi_2$$

Analysis of a DTMC: the "General" Case

Almost general case: every terminal scc is aperiodic

- π_∞ exists.
- $\pi_\infty = \sum_{s \in S} \pi_0(s) \sum_{i \in I} \mathtt{preach}_i[s] \cdot \pi_\infty^i$ where:
  1. S is the set of states,
  2. $\{\mathcal{C}_i\}_{i \in I}$ is the set of terminal scc,
  3. π_∞^i is the steady-state distribution of $\mathcal{C}_i$,
  4. and $\mathtt{preach}_i[s]$ is the probability to reach $\mathcal{C}_i$ starting from s.

Computation of the reachability probability for transient states

- Let T be the set of transient states
  (i.e. not belonging to a terminal scc)
- Let $\mathrm{P}_{T,T}$ be the submatrix of P restricted to transient states
- Let $\mathrm{P}_{T,i}$ be the submatrix of P transitions from T to $\mathcal{C}_i$
- Then $\mathtt{preach}_i = (\sum_{n \in \mathbb{N}} (\mathrm{P}_{T,T})^n) \cdot \mathrm{P}_{T,i} \cdot \mathbf{1} = (Id - \mathrm{P}_{T,T})^{-1} \cdot \mathrm{P}_{T,i} \cdot \mathbf{1}$

Analysis of a DTMC: the "General" Case

Almost general case: every terminal scc is aperiodic

- π_∞ exists.
- $\pi_\infty = \sum_{s \in S} \pi_0(s) \sum_{i \in I} \mathtt{preach}_i[s] \cdot \pi_\infty^i$ where:
  1. S is the set of states,
  2. $\{\mathcal{C}_i\}_{i \in I}$ is the set of terminal scc,
  3. π_∞^i is the steady-state distribution of $\mathcal{C}_i$,
  4. and $\mathtt{preach}_i[s]$ is the probability to reach $\mathcal{C}_i$ starting from s.

Computation of the reachability probability for transient states

- Let T be the set of transient states *(i.e. not belonging to a terminal scc)*
- Let $\mathtt{P}_{T,T}$ be the submatrix of $\mathtt{P}$ restricted to transient states
- Let $\mathtt{P}_{T,i}$ be the submatrix of $\mathtt{P}$ transitions from T to $\mathcal{C}_i$
- Then $\mathtt{preach}_i = \left(\sum_{n \in \mathbb{N}} (\mathtt{P}_{T,T})^n\right) \cdot \mathtt{P}_{T,i} \cdot \mathbf{1} = (Id - \mathtt{P}_{T,T})^{-1} \cdot \mathtt{P}_{T,i} \cdot \mathbf{1}$

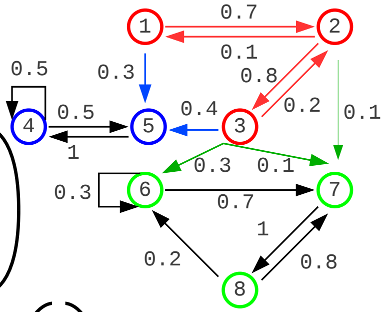
Illustration: SCC and Matrices

$T=\{1,2,3\}, C_1=\{4,5\}, C_2=\{6,7,8\}$

$$P_{T,T}=\begin{pmatrix} 0.0 & 0.7 & 0.0 \\ 0.1 & 0.0 & 0.8 \\ 0.0 & 0.2 & 0.0 \end{pmatrix}$$

$$P_{T,1}.\mathbf{1}=\begin{pmatrix} 0.0 & 0.3 \\ 0.0 & 0.0 \\ 0.0 & 0.4 \end{pmatrix}\begin{pmatrix} 1.0 \\ 1.0 \\ 1.0 \end{pmatrix}=\begin{pmatrix} 0.3 \\ 0.0 \\ 0.4 \end{pmatrix}$$

$$P_{T,2}.\mathbf{1}=\begin{pmatrix} 0.0 & 0.0 & 0.0 \\ 0.0 & 0.1 & 0.0 \\ 0.3 & 0.1 & 0.0 \end{pmatrix}\begin{pmatrix} 1.0 \\ 1.0 \\ 1.0 \end{pmatrix}=\begin{pmatrix} 0.0 \\ 0.1 \\ 0.4 \end{pmatrix}$$



Continuous Time Markov Chain (CTMC)

A CTMC is a stochastic process which fulfills:

- Memoryless state change

$$Pr(S_{n+1} = s_j \mid S_0 = s_{i_0}, ..., S_{n-1} = s_{i_{n-1}}, T_0 < \tau_0, ..., T_n < \tau_n, S_n = s_i)$$
$$= Pr(S_{n+1} = s_j \mid S_n = s_i) \equiv \mathtt{P}[i,j]$$

- Memoryless transition delay

$$Pr(T_n < \tau \mid S_0 = s_{i_0}, ..., S_{n-1} = s_{i_{n-1}}, T_0 < \tau_0, ..., T_{n-1} < \tau_{n-1}, S_n = s_i)$$
$$= Pr(T_n < \tau \mid S_n = s_i) = 1 - e^{-\lambda_i \tau}$$

Notations and properties

- $\mathtt{P}$ defines an *embedded* DTMC (the chain of state changes)
- Let $\pi(\tau)$ the distribution de $X(\tau)$, for δ going to 0 it holds that:

$$\pi(\tau+\delta)(s_i) \approx \pi(\tau)(s_i)(1-\lambda_i\delta) + \sum_j \pi(\tau)(s_j)\lambda_j\delta\mathtt{P}[j,i]$$

- Hence, let $\mathtt{Q}$ *the infinitesimal generator* defined by:
$\mathtt{Q}[i,j] \equiv \lambda_i \mathtt{P}[i,j]$ for $j \neq i$ and $\mathtt{Q}[i,i] \equiv -\sum_{j\neq i} \mathtt{Q}[i,j]$
Then: $\frac{d\pi}{d\tau} = \pi \cdot \mathtt{Q}$

Continuous Time Markov Chain (CTMC)

A CTMC is a stochastic process which fulfills:

- Memoryless state change

$$Pr(S_{n+1} = s_j \mid S_0 = s_{i_0}, ..., S_{n-1} = s_{i_{n-1}}, T_0 < \tau_0, ..., T_n < \tau_n, S_n = s_i)$$
$$= Pr(S_{n+1} = s_j \mid S_n = s_i) \equiv \mathtt{P}[i,j]$$

- Memoryless transition delay

$$Pr(T_n < \tau \mid S_0 = s_{i_0}, ..., S_{n-1} = s_{i_{n-1}}, T_0 < \tau_0, ..., T_{n-1} < \tau_{n-1}, S_n = s_i)$$
$$= Pr(T_n < \tau \mid S_n = s_i) = 1 - e^{-\lambda_i \tau}$$

Notations and properties

- $\mathtt{P}$ defines an *embedded* DTMC (the chain of state changes)
- Let $\pi(\tau)$ the distribution de $X(\tau)$, for δ going to 0 it holds that:
$$\pi(\tau+\delta)(s_i) \approx \pi(\tau)(s_i)(1-\lambda_i\delta) + \sum_j \pi(\tau)(s_j)\lambda_j\delta\mathtt{P}[j,i]$$
- Hence, let $\mathtt{Q}$ *the infinitesimal generator* defined by:
$\mathtt{Q}[i,j] \equiv \lambda_i \mathtt{P}[i,j]$ for $j \neq i$ and $\mathtt{Q}[i,i] \equiv -\sum_{j\neq i} \mathtt{Q}[i,j]$
Then: $\frac{d\pi}{d\tau} = \pi \cdot \mathtt{Q}$

The exponential distribution

Let F be defined by: $F(\tau) = 1 - e^{-\lambda\tau}$
Then F is the *exponential distribution* with *rate* $\lambda > 0$.

The exponential distribution is *memoryless*.
Let X be a random variable with a λ-exponential distribution.

$$\mathbf{Pr}(X > \tau' \mid X > \tau) = \frac{\mathbf{Pr}(X > \tau')}{\mathbf{Pr}(X > \tau)} = \frac{e^{-\lambda\tau'}}{e^{-\lambda\tau}} = e^{-\lambda(\tau' - \tau)} = \mathbf{Pr}(X > \tau' - \tau)$$

The minimum of exponential distributions is an exponential distribution.
Let Y be independent from X with μ-exponential distribution.

$$\mathbf{Pr}(\min(X, Y) > \tau) = e^{-\lambda\tau} e^{-\mu\tau} = e^{-(\lambda+\mu)\tau}$$

The minimal variable is selected proportionally to its rate.

$$\mathbf{Pr}(X < Y) = \int_0^\infty \mathbf{Pr}(Y > \tau) F_X\{d\tau\} = \int_0^\infty e^{-\mu\tau} \lambda e^{-\lambda\tau} d\tau = \frac{\lambda}{\lambda + \mu}$$

Convoluting the exponential distribution

The n^{th} convolution of a distribution F is defined by:

$$F^{n\star} \stackrel{\text{def}}{=} F \star \cdots \star F \qquad (n \text{ times})$$

Let f_n (resp. F_n) be the density (resp. distribution) of the n^{th} convolution of the λ-exponential distribution. Then:

$$f_n(x) = \lambda e^{-\lambda x} \frac{(\lambda x)^{n-1}}{(n-1)!} \text{ and } F_n(x) = 1 - e^{-\lambda x} \sum_{0 \le m < n} \frac{(\lambda x)^m}{m!}$$

Sketch of proof

Recall that: $f_1(x) = \lambda e^{-\lambda x}$.

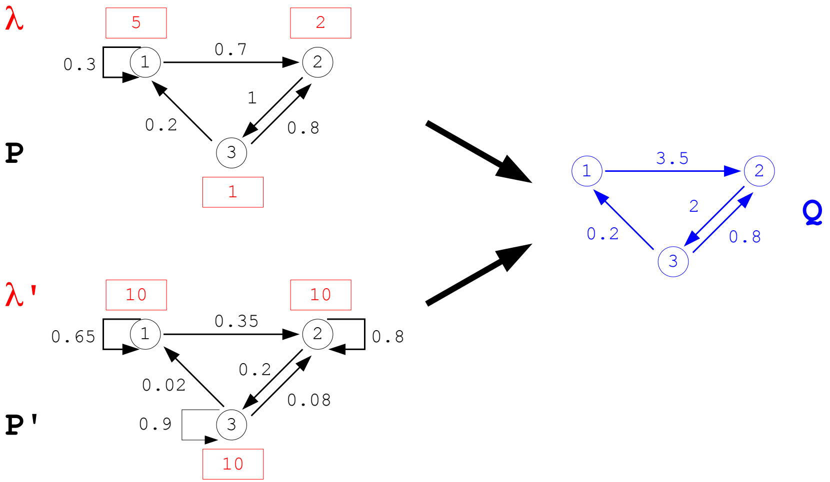
$$f_{n+1}(x) = \int_0^x f_n(x-u) f_1(u) du = \int_0^x \lambda e^{-\lambda(x-u)} \frac{(\lambda(x-u))^{n-1}}{(n-1)!} \lambda e^{-\lambda u} du$$
$$= \lambda e^{-\lambda x} \int_0^x \lambda \frac{(\lambda(x-u))^{n-1}}{(n-1)!} du = \lambda e^{-\lambda x} \frac{(\lambda x)^n}{n!}$$

Deduce F_{n+1} by:

$$\frac{d}{dx}\left(1 - e^{-\lambda x} \sum_{0 \le m \le n} \frac{(\lambda x)^m}{m!}\right) =$$
$$e^{-\lambda x}\left(\lambda \sum_{0 \le m \le n} \frac{(\lambda x)^m}{m!} - \sum_{0 \le m \le n-1} \lambda \frac{(\lambda x)^m}{m!}\right) = f_{n+1}(x)$$

CTMC: Illustration and Uniformization

A CTMC



A *uniform version* of the CTMC *(equivalent w.r.t. the states)*

Analysis of a CTMC

Transient Analysis

- Construction of a uniform version of the CTMC $(\lambda, \mathtt{P})$ such that $\mathtt{P}[i,i] > 0$ for all i.
- Computation by case decomposition w.r.t. the number of transitions:

$$\pi(\tau) = \pi(0) \sum_{n \in \mathbb{N}} (e^{-\lambda\tau}) \frac{\tau^n}{n!} \mathtt{P}^n$$

Steady-state analysis

- The steady-state distribution of visits is given by the steady-state distribution of $(\lambda, \mathtt{P})$ (by construction, the terminal scc are aperiodic) ...
- equal to the steady-state distribution since the sojourn times follow the same distribution.
- A particular case: $\mathtt{P}$ irreducible
the steady-state distribution π is the unique solution of $X \cdot \mathtt{Q} = 0 \wedge X \cdot \mathbf{1} = 1$
where one can omit an arbitrary equation of the first system.

Analysis of a CTMC

Transient Analysis

- Construction of a uniform version of the CTMC $(\lambda, \mathtt{P})$ such that $\mathtt{P}[i,i] > 0$ for all i.
- Computation by case decomposition w.r.t. the number of transitions:

$$\pi(\tau) = \pi(0) \sum_{n \in \mathbb{N}} (e^{-\lambda\tau}) \frac{\tau^n}{n!} \mathtt{P}^n$$

Steady-state analysis

- The steady-state distribution of visits is given by the steady-state distribution of $(\lambda, \mathtt{P})$ (by construction, the terminal scc are aperiodic) ...
- equal to the steady-state distribution since the sojourn times follow the same distribution.
- A particular case: $\mathtt{P}$ irreducible
  the steady-state distribution π is the unique solution of $X \cdot \mathtt{Q} = 0 \wedge X \cdot \mathbf{1} = 1$
  where one can omit an arbitrary equation of the first system.

Plan

1. Stochastic Petri Net

2. Markov Chain

3. Markovian Stochastic Petri Net

4. Generalized Markovian Stochastic Petri Net (GSPN)

5. Product-form Petri Nets

Markovian Stochastic Petri Net

Hypotheses

- The distribution of every transition t_i has a density function $e^{-\lambda_i \tau}$ where the parameter λ_i is called *the rate* of the transition.
- For simplicity reasons, the server policy is single server.

First observations

- **The weights for choice policy are no more required** since equality of two samples has a null probability. *(due to continuity of distributions)*
- The residual delay $d_j - d_i$ of transition t_j knowing that t_i has fired (i.e. d_i is the shortest delay) has the same distribution as the initial delay. **Thus the memory policy is irrelevant.**

Markovian Stochastic Petri Net

Hypotheses

- The distribution of every transition t_i has a density function $e^{-\lambda_i \tau}$ where the parameter λ_i is called *the rate* of the transition.
- For simplicity reasons, the server policy is single server.

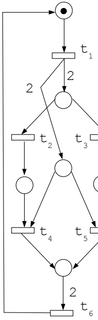
First observations

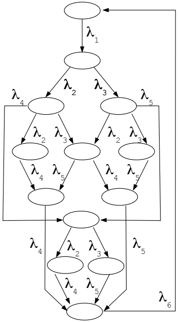
- **The weights for choice policy are no more required** since equality of two samples has a null probability. *(due to continuity of distributions)*
- The residual delay $d_j - d_i$ of transition t_j knowing that t_i has fired (i.e. d_i is the shortest delay) has the same distribution as the initial delay. **Thus the memory policy is irrelevant.**

Markovian Net and Markov Chain

Key observation: given a marking m with $T_m = t_1, \ldots, t_k$

- The sojourn time in m is an exponential distribution with rate $\sum_i \lambda_i$.
- The probability that t_i is the next transition to fire is $\frac{\lambda_i}{(\sum_j \lambda_j)}$.
- **Thus the stochastic process is a CTMC whose states are markings and whose transitions are the transitions of the reachability graph.**





Plan

Generalizing Distributions for Nets

Modelling delays with exponential distributions is **reasonable** when:

- Only mean value information is known about distributions.
- Exponential distributions (or combination of them) are enough to approximate the "real" distributions.

Modelling delays with exponential distributions is **not reasonable** when:

- The distribution of an event is known and is poorly approximable with exponential distributions:

  a time-out of 10 time units
- The delays of the events have different magnitude orders:

  executing an instruction versus performing a database request

In the last case, the 0-Dirac distribution is required.

Generalized Markovian Stochastic Petri Net (GSPN)

Generalized Markovian Stochastic Petri Nets (GSPN) are nets whose:

- *timed transitions* have exponential distributions,
- and *immediate transitions* have 0-Dirac distributions.

Their analysis is based on Markovian Renewal Process,

a generalization of Markov chains.

Markovian Renewal Process

A Markovian Renewal Process (MRP) fulfills:

- a *relative* memoryless property

$$Pr(S_{n+1} = s_j, T_n < \tau \mid S_0 = s_{i_0}, ..., S_{n-1} = s_{i_{n-1}}, T_0 < \tau_0, ..., S_n = s_i) \\ = Pr(S_{n+1} = s_j, T_n < \tau \mid S_n = s_i) \equiv \mathtt{Q}[i, j, \tau]$$

- The embedded chain is defined by: $\mathtt{P}[i, j] = \lim_{\tau \to \infty} \mathtt{Q}[i, j, \tau]$
- The sojourn time `Soj` has a distribution defined by:

$$Pr(\mathtt{Soj}[i] < \tau) = \sum_j \mathtt{Q}[i, j, \tau]$$

Analysis of a MRP

- The steady-state distribution (if there exists) π is deduced from the steady-state distribution of the embedded chain π' by:

$$\pi(s_i) = \frac{\pi'(s_i) E(\mathtt{Soj}[i])}{\sum_j \pi'(s_j) E(\mathtt{Soj}[j])}$$

- Transient analysis is much harder ... but the reachability probabilities only depend on the embedded chain.

Markovian Renewal Process

A Markovian Renewal Process (MRP) fulfills:

- a *relative* memoryless property

$$Pr(S_{n+1} = s_j, T_n < \tau \mid S_0 = s_{i_0}, ..., S_{n-1} = s_{i_{n-1}}, T_0 < \tau_0, ..., S_n = s_i) \\ = Pr(S_{n+1} = s_j, T_n < \tau \mid S_n = s_i) \equiv \mathtt{Q}[i,j,\tau]$$

- The embedded chain is defined by: $\mathtt{P}[i,j] = \lim_{\tau\to\infty} \mathtt{Q}[i,j,\tau]$
- The sojourn time Soj has a distribution defined by:

$$Pr(\mathtt{Soj}[i] < \tau) = \sum_j \mathtt{Q}[i,j,\tau]$$

Analysis of a MRP

- The steady-state distribution (if there exists) π is deduced from the steady-state distribution of the embedded chain π' by:

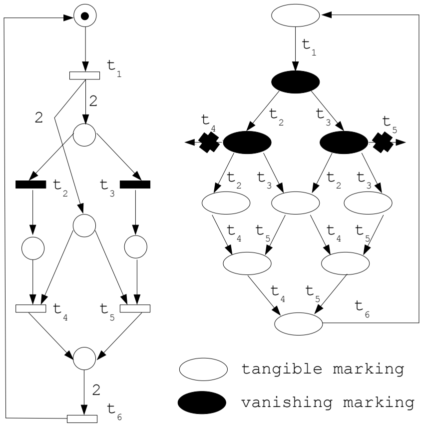
$$\pi(s_i) = \frac{\pi'(s_i)E(\mathtt{Soj}[i])}{\sum_j \pi'(s_j)E(\mathtt{Soj}[j])}$$

- Transient analysis is much harder ... but the reachability probabilities only depend on the embedded chain.

A GSPN is a Markovian Renewal Process

Observations

- Weights are required for immediate transitions.
- The *restricted* reachability graph corresponds to the embedded DTMC.



Steady-State Analysis of a GSPN (1)

Standard method for MRP

- Build the restricted reachability graph equivalent to the embedded DTMC
- Deduce the probability matrix P
- Compute π^* the steady-state distribution of the visits of markings: $\pi^* = \pi^* P$
- Compute π the steady-state distribution of the sojourn in tangible markings:

$$\pi(m) = \frac{\pi^*(m)\mathtt{Soj}(m)}{\sum_{m' \ tangible} \pi^*(m')\mathtt{Soj}(m')}$$

How to eliminate the vanishing markings sooner in the computation?

Steady-State Analysis of a GSPN (2)

An alternative method

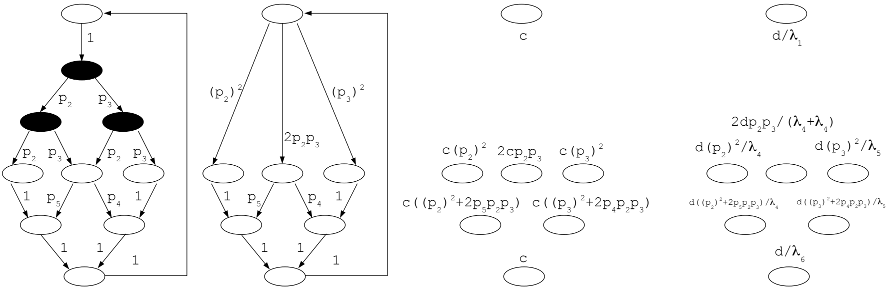
- As before, compute the transition probability matrix P.
- Compute the transition probability matrix P' between tangible markings.
- Compute π'^* the (relative) steady-state distribution of the visits of tangible markings: $\pi'^* = \pi'^* P'$.
- Compute π the steady-state distribution of the sojourn in tangible markings:

$$\pi(m) = \frac{\pi'^*(m)\mathtt{Soj}(m)}{\sum_{m' \ tangible} \pi'^*(m')\mathtt{Soj}(m')}$$

Computation of P'

- Let $P_{X,Y}$ the probability transition matrix from subset X to subset Y.
- Let V (resp. T) be the set of vanishing (resp. tangible) markings.
- $P' = P_{T,T} + P_{T,V}(\sum_{n \in \mathbb{N}} P_{V,V}^n)P_{V,T} = P_{T,T} + P_{T,V}(Id - P_{V,V})^{-1}P_{V,T}$
- Iterative (resp. direct) computations uses the first (resp. second) expression.

Steady-State Analysis: Illustration

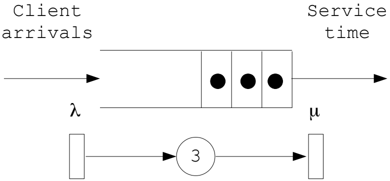


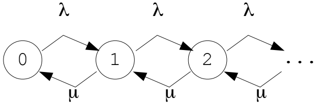
$p_2 = w_2/(w_2+w_3)$ $p_3 = w_3/(w_2+w_3)$
$p_4 = \lambda_4/(\lambda_4+\lambda_5)$ $p_5 = \lambda_5/(\lambda_4+\lambda_5)$

"c" and "d" are normalizing constants

Plan

Steady-State Analysis of a Queue





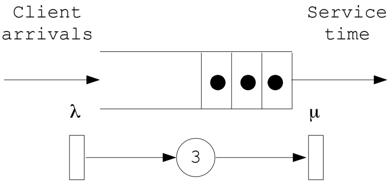
A (Markovian) queue is a CTMC

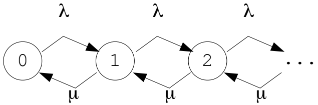
- Interarrival time: exponential distribution with parameter λ
- Service time: exponential distribution with parameter μ

Let $\rho = \frac{\lambda}{\mu}$ be the *utilization*

- The steady-state distribution π_∞ exists iff $\rho < 1$
- The probability of n clients in the queue is $\pi_\infty(n) = \rho^n(1-\rho)$

Steady-State Analysis of a Queue





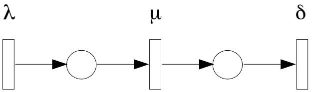
A (Markovian) queue is a CTMC

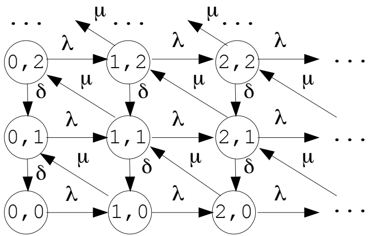
- Interarrival time: exponential distribution with parameter λ
- Service time: exponential distribution with parameter μ

Let $\rho = \frac{\lambda}{\mu}$ be the *utilization*

- The steady-state distribution π_∞ exists iff $\rho < 1$
- The probability of n clients in the queue is $\pi_\infty(n) = \rho^n(1-\rho)$

Analysis of Two Queues in Tandem



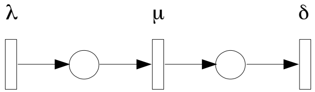


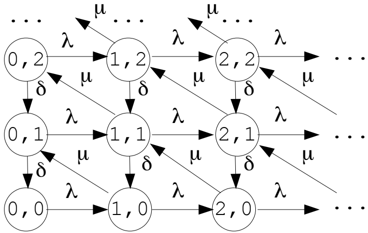
Observation. The associated Markov chain is more complex than the one corresponding to two isolated queues. However ...

Assume $\rho_1 = \frac{\lambda}{\mu} < 1$ and $\rho_2 = \frac{\lambda}{\delta} < 1$

- The steady-state distribution π_∞ exists.
- The probability of n_1 clients in queue 1 and n_2 clients in queue 2 is $\pi_\infty(n_1, n_2) = \rho_1^{n_1}(1-\rho_1)\rho_2^{n_2}(1-\rho_2)$
- It is the product of the steady-state distributions corresponding to two isolated queues.

Analysis of Two Queues in Tandem



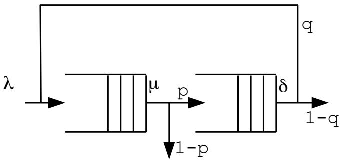


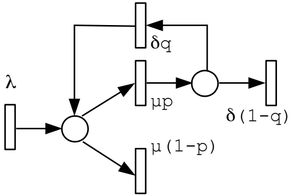
Observation. The associated Markov chain is more complex than the one corresponding to two isolated queues. However ...

Assume $\rho_1 = \frac{\lambda}{\mu} < 1$ and $\rho_2 = \frac{\lambda}{\delta} < 1$

- The steady-state distribution π_∞ exists.
- The probability of n_1 clients in queue 1 and n_2 clients in queue 2 is $\pi_\infty(n_1, n_2) = \rho_1^{n_1}(1-\rho_1)\rho_2^{n_2}(1-\rho_2)$
- It is the product of the steady-state distributions corresponding to two isolated queues.

Analysis of an Open Queuing Network





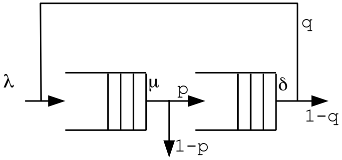
In a steady-state

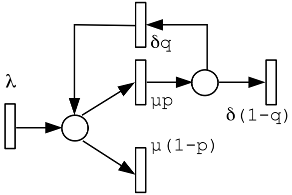
- Define the (input and output) flow through queue 1 (resp. 2) as γ_1 (resp. γ_2).
- Then $\gamma_1 = \lambda + q\gamma_2$ and $\gamma_2 = p\gamma_1$. Thus $\gamma_1 = \frac{\lambda}{1-pq}$ and $\gamma_2 = \frac{p\lambda}{1-pq}$

Assume $\rho_1 = \frac{\gamma_1}{\mu} < 1$ and $\rho_2 = \frac{\gamma_2}{\delta} < 1$

- The steady-state distribution π_∞ exists.
- The probability of n_1 clients in queue 1 and n_2 clients in queue 2 is $\pi_\infty(n_1, n_2) = \rho_1^{n_1}(1-\rho_1)\rho_{n_2}^{n}(1-\rho_2)$
- It is the product of the steady-state distributions corresponding to two isolated queues.

Analysis of an Open Queuing Network





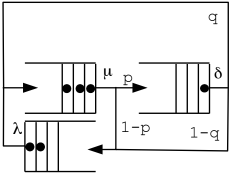
In a steady-state

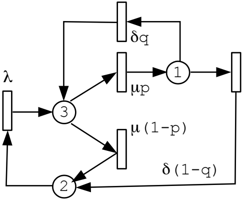
- Define the (input and output) flow through queue 1 (resp. 2) as γ_1 (resp. γ_2).
- Then $\gamma_1 = \lambda + q\gamma_2$ and $\gamma_2 = p\gamma_1$. Thus $\gamma_1 = \frac{\lambda}{1-pq}$ and $\gamma_2 = \frac{p\lambda}{1-pq}$

Assume $\rho_1 = \frac{\gamma_1}{\mu} < 1$ and $\rho_2 = \frac{\gamma_2}{\delta} < 1$

- The steady-state distribution π_∞ exists.
- The probability of n_1 clients in queue 1 and n_2 clients in queue 2 is $\pi_\infty(n_1, n_2) = \rho_1^{n_1}(1-\rho_1)\rho_{n_2}^{n}(1-\rho_2)$
- It is the product of the steady-state distributions corresponding to two isolated queues.

Analysis of a Closed Queuing Network





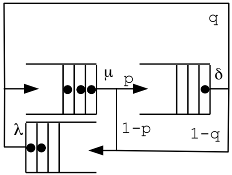
Visit ratios (up to a constant)

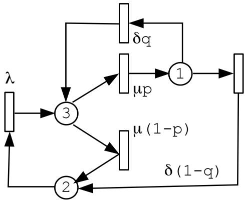
- Define the visit ratio flow of queue i as v_i.
- Then $v_1 = v_3 + qv_2$, $v_2 = pv_1$ and $v_3 = (1-p)v_1 + (1-q)v_2$. Thus $v_1 = 1$, $v_2 = p$ and $v_3 = 1 - pq$.

Define $\rho_1 = \frac{v_1}{\mu}$, $\rho_2 = \frac{v_2}{\delta}$ and $\rho_3 = \frac{v_3}{\lambda}$

- The steady-state probability of n_i clients in queue i is $\pi_\infty(n_1, n_2, n_3) = \frac{1}{G}\rho_1^{n_1}\rho_2^{n_2}\rho_3^{n_3}$ (with $n_1 + n_2 + n_3 = n$)
- where G the normalizing constant can be efficiently computed by dynamic programming.

Analysis of a Closed Queuing Network





Visit ratios (up to a constant)

- Define the visit ratio flow of queue i as v_i.
- Then $v_1 = v_3 + qv_2$, $v_2 = pv_1$ and $v_3 = (1-p)v_1 + (1-q)v_2$.
  Thus $v_1 = 1$, $v_2 = p$ and $v_3 = 1 - pq$.

Define $\rho_1 = \frac{v_1}{\mu}$, $\rho_2 = \frac{v_2}{\delta}$ and $\rho_3 = \frac{v_3}{\lambda}$

- The steady-state probability of n_i clients in queue i is
  $\pi_\infty(n_1, n_2, n_3) = \frac{1}{G}\rho_1^{n_1}\rho_2^{n_2}\rho_3^{n_3}$ (with $n_1 + n_2 + n_3 = n$)
- where G the normalizing constant can be efficiently computed by dynamic programming.

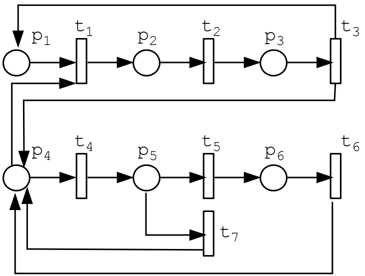
Queuing Networks and Petri Nets

Observations

- A (single client class) queuing network can easily be represented by a Petri net.
- Such a Petri net is a *state machine*: every transition has at most a single input and a single output place.

Can we define a more general subclass of Petri nets with a product form for the steady-state distribution?

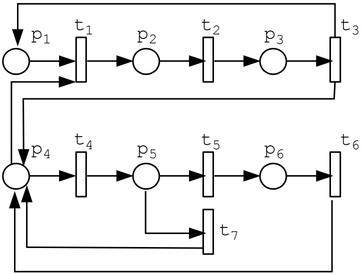
Product Form Stochastic Petri Nets (PFSPN)

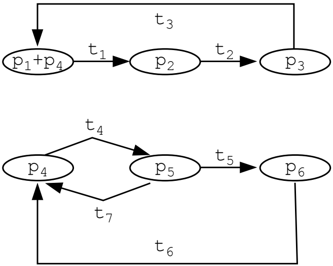


Principles

- Transitions can be partionned into subsets corresponding to several classes of clients with their specific activities
- Places model resources shared between the clients.
- Client states are implicitely represented.

Bags and Transitions in PFSPN



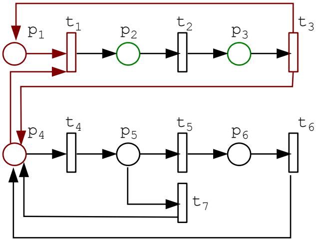


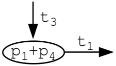
The resource graph

- The vertices are the input and the ouput bags of the transitions.
- Every transition of the net t yields a graph transition ${}^\bullet t \xrightarrow{t} t^\bullet$
- Client classes correspond to the connected components of the graph.

First requirement: The connected components of the graph must be strongly connected.

Witnesses in PFSPN





Vector $-p_2-p_3$ is a witness for bag p_1+p_4:

$(-p_2-p_3) \bullet W(t_3)=1$

$(-p_2-p_3) \bullet W(t_1)=-1$

$(-p_2-p_3) \bullet W(t)=0$ for every other t

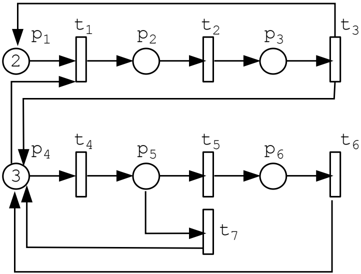
where W is the incidence matrix

Witness for a bag b

- Let $In(b)$ (resp. $Out(b)$) the transitions with input (resp. output) b.
- Let v be a place vector, v is a *witness* for b if:
  - $\forall t \in In(b)\ v \cdot W(t) = -1$ (where $W(t)$ is the incidence of t)
  - $\forall t \in Out(b)\ v \cdot W(t) = 1$
  - $\forall t \notin In(b) \cup Out(b)\ v \cdot W(t) = 0$

Second requirement: Every bag must have a witness.

Steady-State Distributions of PFSPN



The reachability space:

m(p1)+ m(p2)+ m(p3)=2

m(p4)+ m(p5)+ m(p6)=m(p1)+1

Steady-state distribution

- Assume the requirements are fulfilled, with $w(b)$ the witness for bag b.
- Compute the ratio visit of bags $v(b)$ on the resource graph.
- The output rate of a bag b is $\mu(b) = \sum_{t|\bullet t=b} \mu(t)$ with $\mu(t)$ the rate of t.
- Then: $\pi_\infty(m) = \frac{1}{G} \prod_b \left(\frac{v(b)}{\mu(b)}\right)^{w(b)\cdot m}$

Observation. The normalizing constant can be efficiently computed if the reachability space is characterized by linear place invariants.

Some References

M. Ajmone Marsan, G. Balbo, G. Conte, S. Donatelli, G. Franceschinis

Modelling with Generalized Stochastic Petri Nets
Wiley series in parallel computing, John Wiley & Sons, 1995
Freely available on the Web site of GreatSPN

S. Haddad, P. Moreaux

Chapter 7: Stochastic Petri Nets
Petri Nets: Fundamental Models and Applications Wiley pp. 269-302, 2009

S. Haddad, P. Moreaux, M. Sereno, M. Silva

Product-form and stochastic Petri nets: a structural approach.
Performance Evaluation, 59: 313-336, 2005.

S. Haddad, J. Mairesse and H.-T. Nguyen

Synthesis and Analysis of Product-form Petri Nets.
Fundamenta Informaticae 122(1-2), pages 147-172, 2013.