

Chapitre 9

Les réseaux de Petri stochastiques

Serge Haddad , Patrice Moreaux¹

1. Introduction

L'un des intérêts majeurs des réseaux de Petri est de pouvoir combiner l'analyse qualitative (*i.e.* la vérification de propriétés) et quantitative (*i.e.* l'évaluation de performance) [FLO 78, MOL 81, REI 98a, REI 98b]. Par comparaison, les modèles de concurrence tels que les algèbres de processus [HIL 96] n'ont connu que très tardivement des extensions stochastiques et si les premiers résultats sont prometteurs, ils restent en deçà des nombreuses recherches sur l'évaluation de performance des réseaux de Petri stochastiques. De même, les modèles traditionnels d'évaluation de performance comme les réseaux de file d'attente [KLE 75] ne comportent pas de mécanisme de synchronisation et leur adjonction par des moyens ad hoc [FDI 86, DAL 97] n'atteint pas la généralité et la simplicité de la modélisation de la concurrence par les réseaux de Petri.

Les réseaux de Petri stochastiques ont été introduits de manière pragmatique à la fin des années 70, afin de pouvoir bénéficier des méthodes d'évaluation des chaînes de Markov. Cette façon de faire conduit à des résultats immédiats mais ne permet pas d'étudier les aspects sémantiques sous-jacents à la définition des réseaux de Petri stochastiques et ne se généralise pas aisément à des distributions de probabilité quelconques. Aussi dans la présentation des trois

¹LAMSADE, Université Paris Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris cedex 16, France, ({haddad,moreaux}@lamsade.dauphine.fr)

LERI-RESYCOM, Université de Reims Champagne-Ardenne, BP 1039, 51687 Reims cedex 2, France, (patrice.moreaux@univ-reims.fr),

chapitres consacrés aux réseaux de Petri stochastiques, nous procéderons de la façon suivante. Nous nous placerons d'abord au niveau sémantique, c'est à dire au niveau du processus stochastique dont nous étudions les propriétés permettant d'obtenir des méthodes d'analyse. Nous avons choisi ici de mettre en évidence les principes qui caractérisent chaque méthode. Nous occulterons donc les aspects de mise en oeuvre qui relèvent du calcul numérique. En effet, ces aspects ne sont pas spécifiques des réseaux de Petri stochastiques et sont couverts par d'excellents ouvrages [STE 94, BOL 98]. Puis nous présenterons les modèles (et extensions) des réseaux de Petri stochastiques qui engendrent de tels processus stochastiques. Les contraintes qui apparaissent lors de la définition s'expliquent alors intuitivement. Enfin nous développons la méthode d'analyse au niveau des réseaux de Petri stochastiques, ce qui met en évidence les liens avec les méthodes qualitatives et s'inscrit dans la continuité de l'ouvrage.

Ce chapitre débute par la caractérisation des processus stochastiques associés aux systèmes à événements discrets (SED). Nous décrivons d'abord la famille de variables aléatoires de ces processus et donnons leur interprétation en termes de réalisation d'une exécution. Puis nous rappelons brièvement la théorie du renouvellement qui garantit sous des conditions souvent vérifiées, l'existence d'une distribution stationnaire du SED. En effet, nous couvrirons principalement l'étude des systèmes à l'équilibre. Nous énumérons ensuite par ordre de complexité croissante des processus stochastiques typiques pour lesquels la théorie du renouvellement peut être employée en commençant par les chaînes de Markov.

Dans un deuxième temps nous développons les points clefs d'une sémantique stochastique pour les réseaux de Petri. Ceci comprend la spécification d'une variable aléatoire associée au délai de franchissement d'une transition, le critère de choix entre plusieurs transitions franchissables (ou sensibilisées), la prise en compte du degré de franchissement dans le tirage de la variable aléatoire associée à une transition et la mémoire des tirages précédents, une fois le franchissement effectué. Puis nous restreignons la nature des distributions, ce qui conduit aux processus stochastiques étudiés lors de la section précédente. Parmi les familles étudiées, les réseaux de Petri à distributions exponentielle et immédiate, appelés réseaux de Petri stochastiques généralisés, sont considérés comme le modèle standard [Ajm 95]. Nous indiquons, pour toutes ces familles, les méthodes de calcul de la distribution stationnaire à partir du graphe d'accessibilité (lorsque celui-ci est fini).

Les algorithmes de base associés aux différents modèles ont une complexité comparable à la taille du graphe d'accessibilité pour les modèles simples et largement supérieure pour les modèles à lois plus générales. Aussi les techniques plus élaborées peuvent se classer en deux familles : les unes visent des complexités plus réduites que la taille du graphe d'état (*e.g.* en restreignant la catégorie

des réseaux de Petri étudiés) et les autres cherchent à conserver des complexités comparables à la taille du graphe pour des lois plus générales.

La dernière section décrit certaines de ces méthodes en mettant l'accent sur la diversité des approches. Deux autres méthodes majeures et applicables entre autres aux réseaux bien formés seront présentées dans les chapitres suivants. Celles couvertes dans cette section sont :

- la recherche de forme produit, c'est à dire d'une formule qui représente la probabilité stationnaire d'un marquage à partir des paramètres du réseau et du marquage lui-même. Cette méthode illustre l'extension d'une technique d'abord appliquée aux réseaux de files d'attente.
- la recherche de bornes (*e.g.* sur les débits) à partir de la structure du réseau. On remarque ici que la programmation linéaire facilite l'analyse quantitative et qualitative des réseaux de Petri.
- des méthodes approximatives qui tirent profit d'une décomposition du réseau ou d'une transformation du processus stochastique.
- la résolution de réseaux dont une unique place est non bornée. L'application de cette méthode montre que des conditions sur la structure de chaînes de Markov se traduisent simplement en termes de réseaux de Petri.

2. Une sémantique stochastique pour les systèmes à événements discrets

Pour tous les modèles à venir, nous reprendrons les acronymes anglais qui sont les plus répandus dans la littérature associée.

2.1. Le modèle stochastique

Nous supposons connues du lecteur les bases de la théorie des probabilités [FEL 68, FEL 71, TRI 82].

Notations :

- $\Pr(E)$ désigne la probabilité d'un événement E et $\Pr(A | B)$ la probabilité de A sachant B .
- $\mathbb{R}$ (resp. $\mathbb{R}^+$) désigne les réels (resp. les réels non négatifs)
- Une mesure sur $\mathbb{R}$ est la donnée d'une fonction F croissante, continue à droite telle que $\lim_{x \rightarrow -\infty} F(x) = 0$. $F(x)$ représente la mesure de l'intervalle $] -\infty, x]$. La masse de la mesure (finie ou infinie) est $\lim_{x \rightarrow \infty} F(x)$
- Une distribution est une mesure de masse 1

- L'intégration usuelle se note ds où s est la variable d'intégration. L'intégration par rapport à une mesure F se note $F\{ds\}$.
- L'adverbe *presque*, dans des expressions telles que *presque partout* ou *presque sûrement*, signifie pour un ensemble de probabilité 1.

Une exécution d'un SED se caractérise par une suite d'événements $\{e_1, \dots, e_n, \dots\}$ (suite que nous supposons infinie) séparés par des intervalles de temps. Seuls les événements peuvent changer l'état du système.

Formellement, le comportement stochastique d'un SED est déterminé par deux familles de variables aléatoires :

- $X_0, \dots, X_n, \dots$ à valeurs dans l'espace d'états du système $\{s_1, \dots, s_k, \dots\}$ que nous supposons discret. X_0 représente la distribution initiale du système et X_n ($n > 0$) la distribution d'état après le $n^{ième}$ événement. L'occurrence d'un événement ne modifie pas nécessairement l'état du système, par conséquent X_{n+1} peut être égal à X_n .
- $T_0, \dots, T_n, \dots$ à valeurs dans $\mathbb{R}^+$ où T_0 représente l'intervalle de temps avant le premier événement et T_n , ($n > 0$) représente l'intervalle de temps entre le $n^{ième}$ et le $(n+1)^{ième}$ événement. Notons que cet intervalle peut être nul (*e.g.* une suite d'instructions considérées comme instantanées au regard de transactions de base de données avec des entrées/sorties).

A priori, aucune restriction ne devrait être imposée sur ces familles de variables aléatoires. Cependant, pour éviter le caractère pathologique de certaines exécutions, nous excluons la possibilité pour un SED d'exécuter une infinité d'actions en un temps fini. Autrement dit, nous établirons des conditions suffisantes de vérification de l'équation suivante :

$$\sum_{n=0}^{\infty} T_n = \infty \text{ presque sûrement} \quad [1]$$

Cette restriction nous autorise à définir l'état du système à tout instant. Soit $N(t)$, la variable aléatoire définie par :

$$N(t) = \inf\{n \text{ tel que } \sum_{k=0}^n T_k > t\}$$

D'après l'équation [1], $N(t)$ est définie *presque partout*. Comme on peut le voir sur la figure 1, $N(t)$ présente des sauts d'amplitude supérieure à 1. L'état $Y(t)$ du système à l'instant t , est alors simplement $X_{N(t)}$. Il est important de noter que $Y(t)$ n'est pas équivalent au processus stochastique, mais qu'il permet, dans la plupart des cas, de procéder aux analyses standard. Le schéma de la Figure 1 présente une exécution possible du processus et illustre l'interprétation de chacune des variables aléatoires introduites plus haut. Dans cet exemple, le processus est initialement dans l'état s_4 et y reste jusqu'à l'instant t_0 où il passe dans l'état s_6 . À l'instant $t_0 + t_1$, le système visite successivement en un

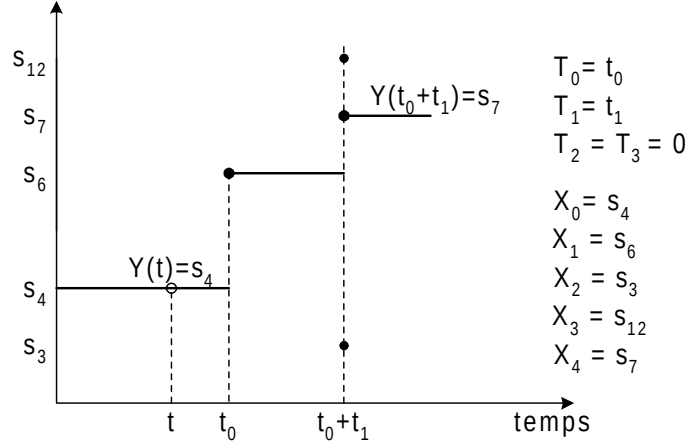


Figure 1: Une réalisation du processus stochastique

temps nul, les états s_3 et s_{12} avant d'atteindre l'état s_7 où il séjourne un certain temps. L'observation $Y(t)$ en temps continu occulte les états évanescents s_3 et s_{12} du processus.

2.2. Analyse par la théorie du renouvellement

L'évaluation de performance d'un SED conduit à deux types d'analyse :

- L'étude du comportement transitoire, c'est à dire l'obtention des mesures d'indices de performances en fonction du temps écoulé depuis l'état initial. Cette étude vise les phases d'initialisation des systèmes et les systèmes à états terminaux. Parmi les domaines d'application, on peut citer l'analyse de fiabilité et de sûreté de fonctionnement [LAP 95, MEY 80, TRI 92].
- L'étude du comportement stationnaire du système. Pour de nombreuses applications, ce qui intéresse le modélisateur est le comportement du système une fois la phase initiale passée, lorsqu'il se stabilise.

Ceci suppose bien entendu qu'un tel comportement stationnaire existe. Ce qui se résume, en notant $\pi(t)$ la distribution de $Y(t)$, par :

$$\lim_{t \rightarrow \infty} \pi(t) = \pi \quad [2]$$

où π est aussi une distribution, appelée *distribution stationnaire*. Dans ce cas, on parlera d'un processus *ergodique*. Une condition suffisante pour ce comportement asymptotique est l'existence d'un phénomène répétitif correspondant à

certaines occurrences d'événements de telle sorte que le processus se comporte de manière identique après chaque occurrence.

Définition 1 *Un processus stochastique est un processus de renouvellement s'il existe une famille de variables aléatoires : $I_1, \dots, I_k, \dots$ (définies presque partout) à valeurs dans $\mathbb{N}$ telles que :*

- $I_k < I_{k+1}$
- $\forall k, k', \{X_{I_k+n}, T_{I_k+n}\}_{n \in \mathbb{N}}$ et $\{X_{I_{k'}+n}, T_{I_{k'}+n}\}_{n \in \mathbb{N}}$ sont des répliques probabilistes d'un processus fixe

Un processus de renouvellement est entièrement déterminé par son comportement entre deux instants de renouvellement. Appelons :

- F la distribution du temps entre deux instants de renouvellement
- d la moyenne, supposée finie, de cette distribution
- $p_k(t)$ la probabilité que, t unités de temps après un instant de renouvellement, il n'y ait pas eu un nouvel instant de renouvellement et que le processus soit dans l'état s_k

Remarquons que $1 - F(t) = \sum_k p_k(t)$. La famille $\{p_k\}_{k \in \mathbb{N}}$ détermine donc le processus stochastique.

Nous devons distinguer deux cas selon la nature de la distribution de F car certaines distributions conduisent à des comportements périodiques. Prenons un SED élémentaire (un processus semi-markovien à durées déterministes, voir plus bas), visitant trois états s_1, s_2, s_3 . Nous considérons une entrée dans l'état s_1 comme un instant de renouvellement. Le SED reste 1s dans l'état s_1 puis passe dans l'état s_2 où il reste durant 2s. Avec une probabilité 1/2, il retourne dans l'état s_1 , ou passe dans l'état s_3 où il reste 3s, avant de revenir en s_1 .

Supposons que le processus démarre dans l'état s_1 . Alors toutes les $3n$ s, le processus est soit en s_1 , soit en s_3 et toutes les $(3n+1)$ s le processus est soit en s_2 , soit en s_3 . Il n'y a donc pas de distribution stationnaire. Ces comportements périodiques sont engendrés par les distributions arithmétiques.

Définition 2 *Une distribution est arithmétique si elle est concentrée sur les points $\{n \cdot \tau\}_{n \in \mathbb{N}}$. La période d'une distribution arithmétique est le plus grand τ vérifiant cette propriété.*

Nous pouvons énoncer alors le résultat principal sur l'existence d'un comportement stationnaire.

Théorème 3 ([FEL 68, FEL 71]) *Soit un processus de renouvellement défini par $\{p_k\}_{k \in \mathbb{N}}$. Si F n'est pas arithmétique alors,*

$$\lim_{t \rightarrow \infty} \pi(t)[s_k] = \frac{1}{d} \cdot \int_0^\infty p_k(t) dt$$

Si F est arithmétique de période τ et si le processus démarre à un instant de renouvellement,

$$\lim_{n \rightarrow \infty} \boldsymbol{\pi}(t + n.\tau)[s_k] = \frac{\tau}{d} \cdot \sum_{i=0}^{\infty} p_k(t + i.\tau)$$

Si les deux formules donnent l'illusion d'un caractère opérationnel, il n'en est rien et la probabilité stationnaire est généralement calculée par d'autres procédures comme nous allons le voir sur les exemples suivants qui sont typiques des processus de renouvellement associés aux SED. Ces formules sont généralisables à tout indice de performance qui est indépendant du comportement du processus avant le dernier instant de renouvellement (*e.g.* le temps écoulé depuis le dernier instant de renouvellement).

2.3. Chaînes de Markov à temps discret

2.3.1. Présentation

Une chaîne de Markov à temps discret (CMTD) (*Discrete Time Markov Chain*, DTMC) possède les caractéristiques suivantes :

- L'intervalle de temps entre les instants T_n est une constante de valeur 1
- L'état suivant un état atteint ne dépend que de cet état et les probabilités de transition restent constantes² au cours du temps :

$$\Pr(X_{n+1} = s_j \mid X_0 = s_{i_0}, \dots, X_n = s_i) =$$

$$\Pr(X_{n+1} = s_j \mid X_n = s_i) = p_{ij} = \mathbf{P}[i, j]$$

Nous utiliserons indifféremment les deux notations pour les transitions d'état.

Le processus se caractérise par sa distribution initiale $\boldsymbol{\pi}_0$ et la matrice $\mathbf{P}$. Soit $\boldsymbol{\pi}_n$ la distribution de X_n , alors $\boldsymbol{\pi}_n = \boldsymbol{\pi}_0.\mathbf{P}^n$

2.3.2. Conditions d'existence d'une distribution stationnaire

Il est clair que chaque entrée dans un état donné constitue un instant de renouvellement. Il reste cependant deux conditions à vérifier :

- Il doit y avoir *presque sûrement* une infinité d'instant de renouvellement.
- Le temps moyen entre deux instants de renouvellement doit être fini.

²d'où le terme de chaîne *homogène* utilisé dans les études sur les chaînes de Markov en toute généralité

Ceci nous conduit à une classification des états :

- Un état est *transitoire* si la probabilité d'y revenir est inférieure à 1. Un tel état ne peut constituer un processus de renouvellement car le nombre de retours est *presque sûrement* fini. De manière évidente, sa probabilité d'occurrence tend vers 0. Un état est appelé *récurrent* s'il n'est pas transitoire.
- Un état est *récurrent nul* si la durée moyenne du retour à cet état est infinie. Il ne peut assurer l'existence d'une distribution stationnaire. Intuitivement, une fois atteint, cet état apparaîtra à des intervalles dont la durée moyenne tendra vers l'infini et par conséquent sa probabilité d'occurrence tendra aussi vers 0. Ce raisonnement intuitif est mathématiquement justifié.
- Un état est *récurrent non nul* si la durée moyenne du retour à cet état est fini. Il suffit alors, pour qu'un tel état puisse constituer le support d'un processus de renouvellement, que cet état soit atteignable *presque sûrement* à partir de la distribution initiale.

Pour approfondir ce point, considérons le graphe (éventuellement infini) construit de la manière suivante :

- L'ensemble des noeuds est un sous-ensemble des états
- Il y a un arc de s_i à s_j si $p_{ij} > 0$.

On restreint les noeuds aux états accessibles par les états dont la probabilité d'occurrence dans la distribution initiale est non nulle.

Étudions les composantes fortement connexes (c.f.c) de ce graphe. Si une c.f.c a un arc sortant, alors nécessairement, les états de cette c.f.c sont transitoires. S'il existe deux c.f.c puits (*i.e.* sans arcs sortants), la probabilité d'atteindre chacune d'entre elles est non nulle, aussi aucune ne peut être atteinte avec certitude. En conséquence, l'existence d'une distribution stationnaire requiert l'existence d'une unique c.f.c puits accessible *presque sûrement*.

On appelle chaîne irréductible, une c.f.c puits. Dans une chaîne irréductible tous les états sont de même type. Examinons la chaîne irréductible définie par :

$$\forall i, p_{i i+1} = 1 - e_i \quad \text{et} \quad p_{i 1} = e_i, \quad \text{avec } 0 < e_i < 1$$

Alors $\Pr(\text{de ne pas revenir en } s_1) = \prod_{i=1}^{\infty} (1 - e_i)$. Un passage au logarithme montre que cette probabilité est non nulle si et seulement si $\sum e_i$ converge. Supposons maintenant que les états sont récurrents; la durée moyenne d'un retour en s_1 est égale à $1 + \sum_{k=1}^{\infty} \prod_{i=1}^k e_i$. Aussi les états sont récurrents nuls si cette somme diverge et récurrents non nuls si elle converge.

Des critères de classification généraux existent (voir plus bas). Dans un graphe fini, l'existence d'une distribution stationnaire est assurée dès lors qu'on a une seule c.f.c puits.

Dans une chaîne de Markov à temps discret, la distribution du retour à un état est forcément arithmétique et sa période est un multiple de 1. Cependant puisque le séjour dans un état dure au moins 1s, si la période est de 1s le cas arithmétique du théorème 3 se ramène au cas non arithmétique et nous sommes assurés d'une distribution stationnaire. On parle alors d'états *ergodiques* et d'une chaîne *apériodique*. Si la période (k) est supérieure à 1, on peut partitionner les états en sous-ensembles $S_0, S_1, \dots, S_{k-1}$ tels que des états de S_i on accède aux états de $S_{(i+1) \bmod k}$. Si on considère les changements d'état toutes les k s (matrice de transition $\mathbf{P}^k$) on retrouve k chaînes indépendantes sur les états S_i de période 1.

2.3.3. Calcul de la distribution stationnaire

Une fois l'existence de la distribution assurée, le calcul est relativement facile. En effet, on a $\pi_{n+1} = \pi_n \cdot \mathbf{P}$. En passant à la limite (justifiée), on obtient $\pi = \pi \cdot \mathbf{P}$. Ce qui y est plus intéressant est que si la chaîne est apériodique, alors π est la seule distribution solution de :

$$\mathbf{X} = \mathbf{X} \cdot \mathbf{P} \quad [3]$$

et que l'existence d'une distribution solution assure que la chaîne irréductible est ergodique.

Dans le cas fini, pour résoudre l'équation [3], on peut procéder à un calcul direct en complétant par l'équation de normalisation $\mathbf{X} \cdot \mathbf{1}^T = 1$ où $\mathbf{1}^T$ désigne le vecteur colonne composé de 1. Mais les calculs itératifs sont plus intéressants, le plus simple consistant à itérer $\mathbf{X}_{n+1} \leftarrow \mathbf{X}_n \cdot \mathbf{P}$ [STE 94].

2.4. Chaînes de Markov à temps continu

2.4.1. Présentation

Une chaîne de Markov à temps continu (CMTC) (*Continuous Time Markov Chain*) a les caractéristiques suivantes :

- L'intervalle de temps entre les instants T_n est une variable aléatoire exponentielle négative dont le taux ne dépend que de l'état X_n . Autrement dit

$$\begin{aligned} \Pr(T_n \leq t \mid X_0 = s_{i_0}, \dots, X_n = s_i, T_0 \leq t_0, \dots, T_{n-1} \leq t_{n-1}) = \\ \Pr(T_n \leq t \mid X_n = s_i) = 1 - e^{-\lambda_i \cdot t} \end{aligned}$$

- L'état suivant un état courant ne dépend que de cet état et les probabilités de transition restent constantes³ au cours du temps :

$$\begin{aligned}\Pr(X_{n+1} = s_j \mid X_0 = s_{i_0}, \dots, X_n = s_i, T_0 \leq t_0, \dots, T_{n-1} \leq t_{n-1}) = \\ \Pr(X_{n+1} = s_j \mid X_n = s_i) = p_{ij} = \mathbf{P}[i, j]\end{aligned}$$

Dans les chaînes de Markov à temps continu, en raison de l'absence de mémoire de la loi exponentielle, l'évolution du SED à tout instant est uniquement conditionnée par son état courant.

Contrairement aux chaînes à temps discret, le processus stochastique peut présenter le caractère pathologique décrit au début de la section. On est assuré d'un comportement normal si, par exemple, il existe une borne supérieure à l'ensemble des λ_i ou encore si la chaîne discrète définie par la matrice $\mathbf{P}$, appelée *chaîne incluse*, est irréductible et récurrente.

Le processus se caractérise par sa distribution initiale $\boldsymbol{\pi}(0)$, la matrice $\mathbf{P}$ et les λ_i . Appelons $\boldsymbol{\pi}(t)$ la distribution de Y_t et $\pi_k(t) = \boldsymbol{\pi}(t)[s_k]$. Si δ est petit, entre t et $t + \delta$ la probabilité de l'occurrence de plus d'un événement est négligeable et la probabilité d'occurrence d'un changement d'état de k à k' est approximativement égale à $\lambda_k \cdot \delta \cdot p_{kk'}$

$$\pi_k(t + \delta) \approx \pi_k(t) \cdot (1 - \lambda_k \cdot \delta) + \sum_{k' \neq k} \pi_{k'}(t) \cdot \lambda_{k'} \cdot \delta \cdot p_{k'k}$$

D'où

$$\frac{\pi_k(t + \delta) - \pi_k(t)}{\delta} \approx \pi_k(t) \cdot (-\lambda_k) + \sum_{k' \neq k} \pi_{k'}(t) \cdot \lambda_{k'} \cdot p_{k'k}$$

Et finalement :

$$\frac{d\pi_k}{dt} = \pi_k(t) \cdot (-\lambda_k) + \sum_{k' \neq k} \pi_{k'}(t) \cdot \lambda_{k'} \cdot p_{k'k}$$

Définissons la matrice $\mathbf{Q}$ par : $q_{kk'} = \lambda_k \cdot p_{kk'}$ pour $k \neq k'$ et $q_{kk} = -\lambda_k (= -\sum_{k' \neq k} q_{kk'})$. Alors l'équation précédente se réécrit :

$$\frac{d\boldsymbol{\pi}}{dt} = \boldsymbol{\pi} \cdot \mathbf{Q} \quad [4]$$

Par un passage à la limite analogue, on obtient le comportement transitoire du processus :

$$\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0) \cdot \sum_{n=0}^{\infty} \frac{t^n}{n!} \cdot \mathbf{Q}^n = \boldsymbol{\pi}(0) \cdot e^{t \cdot \mathbf{Q}} \quad [5]$$

où la deuxième égalité est une définition. On appelle $\mathbf{Q}$ le *générateur infiniésimal* du processus.

³ici aussi, on parle de chaîne *homogène*

2.4.2. Conditions d'existence et calcul d'une distribution stationnaire

Supposons que $\pi(t)$ tende vers une distribution stationnaire. Il est raisonnable d'envisager que $\frac{d\pi}{dt}$ tende vers 0. Donc l'équation [4] devient $\pi \cdot \mathbf{Q} = 0$. De fait, l'existence d'une distribution solution de l'équation :

$$\mathbf{X} \cdot \mathbf{Q} = 0 \quad \text{et} \quad \mathbf{X} \cdot \mathbf{1}^T = 1 \quad [6]$$

est une condition nécessaire et suffisante et ici aussi cette équation admet au plus une solution (si la chaîne incluse est irréductible). Le calcul s'effectue de façon similaire aux chaînes à temps discret.

2.5. Processus semi-markoviens

2.5.1. Présentation

Nous nous restreindrons ici à une notion de processus semi-markovien moins générale que la définition habituelle, et ceci pour deux raisons. D'une part, cette définition rend possible un calcul simplifié des distributions stationnaires ; d'autre part, la section suivante présente une famille de processus encore plus générale que la famille des processus semi-markoviens. Un processus semi-markovien est une extension des chaînes de Markov à temps continu où les temps de séjour dans les états ont une distribution quelconque. Ce processus possède les caractéristiques suivantes :

- L'intervalle de temps entre les instants T_n est une variable aléatoire qui ne dépend que de l'état X_n . Autrement dit :

$$\Pr(T_n \leq t \mid X_0 = s_{i_0}, \dots, X_n = s_i, T_0 \leq t_0, \dots, T_{n-1} \leq t_{n-1}) =$$

$$\Pr(T_n \leq t \mid X_n = s_i) = \Pr(D_i \leq t)$$

où D_i est une variable aléatoire quelconque de distribution à espérance finie, notée d_i .

- L'état suivant un état courant ne dépend que de cet état et les probabilités de transition restent constantes au cours du temps :

$$\Pr(X_{n+1} = s_j \mid X_0 = s_{i_0}, \dots, X_n = s_i, T_0 \leq t_0, \dots, T_{n-1} \leq t_{n-1}) =$$

$$\Pr(X_{n+1} = s_j \mid X_n = s_i) = p_{ij} = \mathbf{P}[i, j]$$

On remarquera, qu'ici aussi, les changements d'états (X_n) constituent une CMTD, incluse dans le processus.

2.5.2. Conditions d'existence et calcul d'une distribution stationnaire

On se limitera ici à spécifier une condition suffisante d'existence d'une distribution stationnaire qui couvre les cas les plus fréquents de stationnarité. Nous commençons par supposer que la chaîne incluse est irréductible avec une distribution solution de $\mathbf{X.P} = \mathbf{X}$ et que l'une des distributions D_i est non arithmétique.

Ici aussi les entrées dans un état (s_i) peuvent constituer un processus de renouvellement. Soit un état donné; le fait qu'il se répète indéfiniment ne dépend que des probabilités de transition p_{ij} et est assuré par notre première hypothèse. Par contre, l'espérance du temps de retour s'analyse différemment. En effet, chaque visite en s_i donne lieu à un séjour d'une durée moyenne d_i . Il se peut alors que, bien que le nombre moyen de visites avant un retour soit fini, la moyenne du temps de retour soit infini. Appelons π' ($\pi'_k = \pi'[s_k]$) la distribution solution de l'équation [3]. Alors le nombre moyen de visites à s_k entre deux visites à s_i est $\frac{\pi'_k}{\pi'_i}$. D'où un temps moyen de retour en s_i égal à :

$$d_i + \sum_{k \neq i} d_k \cdot \frac{\pi'_k}{\pi'_i} = \frac{1}{\pi'_i} \cdot \sum_k d_k \cdot \pi'_k$$

Autrement dit, l'existence d'une distribution stationnaire est assurée si $\sum_k d_k \cdot \pi'_k$ est finie. Ayant fait l'hypothèse que D_i est non arithmétique, on en déduit facilement que la distribution de retour est aussi non arithmétique.

Avec le même raisonnement, on voit que le ratio $\frac{\pi_k}{\pi_i}$ doit correspondre au temps de séjour moyen en s_k entre deux retours en s_i divisé par le temps de séjour moyen en s_i soit :

$$\frac{\pi_k}{\pi_i} = \frac{d_k \cdot \frac{\pi'_k}{\pi'_i}}{d_i} = \frac{d_k \cdot \pi'_k}{d_i \cdot \pi'_i}$$

Ce qui conduit à la distribution stationnaire :

$$\pi_k = \frac{\pi'_k \cdot d_k}{\sum_{k'} \pi'_{k'} \cdot d_{k'}} \quad [7]$$

Notons que cette manière de procéder autorise certaines distributions D_i à être concentrées en 0 (voir la section 3.3).

2.6. Processus de Markov régénératifs

2.6.1. Présentation

Un processus de Markov régénératif (ou processus semi-régénératif) possède un sous-ensemble d'états, appelés états de régénération parce que l'entrée dans chacun de ces états constitue un processus de renouvellement. On appellera S' ce sous-ensemble d'états. Un tel processus est entièrement déterminé par son comportement entre deux entrées dans des états de régénération. Formellement, on définit pour tous $k, k' \in S'$ et tout $i \in S$ les quantités suivantes :

- $F_k(t)$ représente la distribution du temps entre une entrée dans s_k et une entrée dans un nouvel état de régénération
- d_k la moyenne de cette distribution supposée finie
- $f_{ki}(t)$ la probabilité que, t unités de temps après une entrée dans s_k , il n'y ait pas eu une entrée dans un nouvel état de régénération et que le processus soit dans l'état s_i
- $G_{kk'}(t)$ la probabilité qu'après une entrée dans s_k , le processus ait atteint comme nouvel état de régénération $s_{k'}$ en un temps $\leq t$. $G_{kk'}$ représente une mesure de masse ≤ 1 .

2.6.2. Conditions d'existence et calcul d'une distribution stationnaire

Nous décrivons simultanément des conditions suffisantes d'existence d'une distribution stationnaire et son calcul. Nous supposons d'abord que la probabilité d'atteindre dans le futur un nouvel état de régénération à partir d'un état de régénération quelconque est toujours égale à 1, ce qui est équivalent à dire que F_k est une distribution de probabilité.

On se ramène à une chaîne de Markov incluse représentant les visites aux états de régénération de matrice $\mathbf{P}$. Cette matrice se calcule par :

$$p_{kk'} = G_{kk'}(\infty) = \int_0^\infty G_{kk'}\{dt\}$$

Nous supposons que cette chaîne est ergodique et nous notons la distribution solution comme précédemment π' . Il s'agit maintenant de calculer le temps moyen de séjour d'un état s_i entre l'entrée dans s_k et l'entrée dans un autre point de régénération (noté d_{ki}) :

$$d_{ki} = \int_0^\infty f_{ki}(t)dt$$

La distribution stationnaire (π) s'en déduit en pondérant par les probabilités des visites aux états de régénération.

$$\pi_i = \frac{\sum_{k \in S'} \pi'_k \cdot d_{ki}}{\sum_{k \in S'} \pi'_k \cdot d_k}$$

La principale difficulté tient à la détermination des f_{ki} et des $G_{kk'}$. Dans certains cas, ceci se ramène à l'étude d'une chaîne de Markov en régime transitoire (voir la section 3.4).

3. Les réseaux de Petri stochastiques

3.1. Les réseaux de Petri à lois générales

L'aspect stochastique des réseaux de Petri est introduit en considérant qu'une transition possède un délai de franchissement variable résultant d'un tirage aléatoire dans une distribution donnée (à valeurs dans $\mathbb{R}^+$). Les différentes familles de réseaux de Petri stochastiques sont définies en restreignant la nature de ces distributions. Mais pour l'immédiat, nous ne faisons aucune hypothèse sur ces distributions. La définition de ces distributions ne suffit pas à caractériser le processus stochastique. Nous allons successivement étudier les problèmes liés à cette caractérisation.

Remarque : La plupart des paramètres de ce processus peuvent être rendus dépendants du marquage courant. Par souci de simplification, nous ne le mentionnerons pas dans la discussion qui suit.

3.1.1. Politique de choix

Etant donné le marquage initial, il s'agit de déterminer la prochaine transition à franchir parmi les transitions franchissables. Deux possibilités s'offrent à nous :

- un choix probabiliste suivant une distribution associée au sous-ensemble des transitions franchissables. On parle alors de *présélection* car ce choix a lieu avant le tirage aléatoire du délai.
- à l'inverse, un tirage aléatoire indépendant pour chacun des délais suivi du choix du plus court délai. En cas de délais égaux, on effectue également un choix probabiliste, appelé *post-sélection*.

La deuxième solution est systématiquement retenue car d'une part elle correspond à une modélisation plus naturelle et d'autre part parce qu'à l'aide de transitions immédiates (voir la section 3.3) la présélection est simulable par la

post-sélection. Notons qu'excepté si les distributions des délais sont continues, il importe de spécifier les distributions des sélections.

3.1.2. Politique de service

Si une transition a un degré de franchissement $e > 1$, on peut considérer que le marquage *fournit* e clients à la transition vue comme un serveur. Aussi lors du tirage aléatoire du délai, trois options sont possibles selon l'événement modélisé :

- un seul tirage est effectué, la transition n'effectue qu'un service à la fois (mode *single-server*)
- e tirages sont effectués, la transition est un serveur *réentrant* (mode *infinite-server*)
- $\min(e, \deg(t))$ sont effectués, la transition pouvant assurer au plus $\deg(t)$ services simultanés ; ce cas généralise les deux précédents (avec $\deg(t) = 1$ ou ∞) (mode *multiple-server*). Le modélisateur doit alors spécifier $\deg(t)$ pour chaque transition.

3.1.3. Politique de mémoire

Une fois la transition t franchie, quelle est l'influence d'un tirage non retenu d'une transition t' sur ses franchissements futurs ?

La première possibilité consiste à oublier le tirage effectué. Si la transition t' reste franchissable, elle donne lieu à un nouveau tirage aléatoire (mode *resampling memory*). Avec une telle sémantique, t pourrait modéliser un phénomène de panne d'un service spécifié par t' .

La deuxième possibilité consiste à mémoriser le tirage décrétement du tirage retenu, uniquement si t' reste franchissable (mode *enabling memory PRD* (*Preemptive Repeat Different*)). Si t' devient infranchissable, ce mécanisme modélise un *time-out* (t') désarmé par t .

La troisième possibilité est équivalente pour une transition encore franchissable mais conserve le tirage inchangé si t' devient infranchissable. Ce tirage sera utilisé à nouveau lorsque t' redevient franchissable (mode *enabling memory PRI* (*Preemptive Repeat Identical*)). Une transition t' devenue infranchissable pourrait modéliser un travail avorté par t pour être recommencé plus tard.

La quatrième possibilité consiste à mémoriser le tirage décrétement du tirage retenu. Une transition t' devenue infranchissable pourrait modéliser un travail suspendu par t (mode *age memory* aussi appelé *PRS* (*Preemptive ReSume*)).

Pour compléter cette politique, nous devons prendre en compte le cas des transitions réentrantes (*multiple-server*), ce qui nécessite parfois de choisir quels tirages conserver, suspendre ou oublier. La solution la plus simple est une politique *FIFO* des tirages. Le dernier tirage effectué est le premier suspendu ou oublié. D'autres politiques (comme suspendre ou oublier le client le moins engagé) ne sont pas nécessairement compatibles avec certaines méthodes d'analyse.

Il est clair qu'une fois définies ces trois politiques, le processus stochastique est déterminé sans ambiguïté. Nous allons maintenant nous focaliser sur la nature des distributions des délais.

3.2. Les réseaux de Petri à lois exponentielles

Dans le modèle de base [FLO 85, MOL 81] appelé réseau de Petri stochastique (RdPS) (*Stochastic Petri Net*, SPN), chaque transition (t) a un délai exponentiel de taux $\mathbf{w}[t]$ (on notera $w_k = \mathbf{w}[t_k]$).

Examinons le processus stochastique engendré par un RdPS en mode *single-server*. Soit m un marquage donné, $t_1, \dots, t_k$ les transitions franchissables à partir de m . On vérifie que :

- la durée de séjour est une loi exponentielle de taux $w_1 + \dots + w_k$
- la probabilité de franchir t_i avant les autres transitions est égale à $\frac{w_i}{w_1 + \dots + w_k}$ et qu'elle est indépendante du temps de séjour dans le marquage
- la distribution du temps résiduel du délai de franchissement de t_i si t_j est franchie est égale à la distribution initiale (loi sans mémoire)

Autrement dit, seul le nouveau marquage détermine le comportement futur du processus stochastique. On a donc affaire à une chaîne de Markov à temps continu, isomorphe au graphe d'accessibilité du RdP, dont tous les paramètres sont donnés par les états (*i.e.* les marquages). Ce raisonnement se généralise aux autres modes de fonctionnement.

Si le graphe est fini, alors la formule [5] nous donne le comportement transitoire du RdPS et si de plus il possède une seule c.f.c, alors la résolution de l'équation [6] nous fournit la distribution stationnaire du RdPS.

A partir de la distribution stationnaire, d'autres indices de performance peuvent être calculés comme le débit moyen (nombre de franchissements par unités de temps) des transitions donné par :

$$\bar{\chi}_k = \sum_{m \text{ accessible}} \pi_m \cdot \text{services}(m, t_k) \cdot w_k \quad [8]$$

où $\text{services}(m, t_k)$ indique le nombre de clients dans l'état m servis par la transition t_k ; ce nombre dépend du degré de franchissement et de la politique de service de la transition.

3.3. Les réseaux de Petri stochastiques généralisés

Modéliser un algorithme ou un protocole nécessite de représenter des choix, des boucles et d'autres structures de contrôle. Ces actions sont des opérations logiques et ont une durée négligeable au regard d'une transmission de données par exemple. Les modéliser par une loi exponentielle à taux très élevé n'est pas satisfaisant car, d'une part le choix du taux est arbitraire et d'autre part les calculs numériques souffrent de valeurs d'amplitudes très différentes. Pour pallier cette difficulté, les transitions immédiates (*i.e.* avec une distribution concentrée en 0) ont été introduites. Dans ce nouveau modèle [Ajm 84], appelé réseau de Petri stochastique généralisé (RdPSG) (*Generalized Stochastic Petri Net*, GSPN), les marquages se décomposent en deux catégories : les marquages tangibles dans lesquels aucune transition immédiate n'est franchissable et les marquages évanescents.

Examinons le processus stochastique engendré par un RdPSG à partir de m , un marquage donné. Si m est tangible, le processus est identique au RdPS. Examinons le cas d'un marquage évanescent ; les transitions franchissables comportent au moins une transition immédiate. *Presque sûrement* le tirage des transitions exponentielles sera > 0 . Le choix de la transition se fait par une post-sélection entre les transitions immédiates. Puisque la durée des transitions immédiates est nulle et que les lois des autres transitions sont sans mémoire, les délais résiduels sont identiques au délais initiaux et l'état du processus ne dépend que du nouveau marquage.

Nous avons donc affaire à un processus semi-markovien dont les séjours dans les marquages tangibles suivent une loi exponentielle et les durées de séjour dans les marquages évanescents sont nulles. Les probabilités de transition (matrice $\mathbf{P}$) sont obtenues selon le cas soit à partir des taux, soit à partir des paramètres de post-sélection.

L'analyse de la section 2.5.2 s'applique naturellement. Cependant, dans ce cas précis, une amélioration est possible. Remarquons d'abord que dans la distribution stationnaire (voir l'équation [7]), les marquages évanescents ont une probabilité nulle d'occurrence. Aussi il est tentant de les éliminer avant la résolution de la chaîne incluse. Pour ce faire, on considère le processus comme un processus de Markov régénératif dont les états de régénération sont les marquages tangibles. Des paramètres de ce processus, seules les probabilités de

transition entre états de régénération ne sont pas directement accessibles. Pour les obtenir, nous décomposons la matrice $\mathbf{P}$ en sous-matrices :

- $\mathbf{P}_{VV}$ transitions entre marquages évanescents
- $\mathbf{P}_{TT}$ transitions entre marquages tangibles
- $\mathbf{P}_{VT}$ transitions des marquages évanescents vers les marquages tangibles
- $\mathbf{P}_{TV}$ transitions des marquages tangibles vers les marquages évanescents

En raisonnant sur le nombre de marquages évanescents rencontrés, en passant d'un marquage tangible à un autre marquage tangible, on vérifie que la nouvelle matrice de transition ($\mathbf{P}'$) est donnée par :

$$\mathbf{P}' = \mathbf{P}_{TT} + \sum_{n=0}^{\infty} \mathbf{P}_{TV} \cdot (\mathbf{P}_{VV})^n \cdot \mathbf{P}_{VT} = \mathbf{P}_{TT} + \mathbf{P}_{TV} \cdot (\mathbf{Id}_{VV} - \mathbf{P}_{VV})^{-1} \cdot \mathbf{P}_{VT}$$

où $\mathbf{Id}_{VV}$ est la matrice identité sur les marquages évanescents.

Si $\mathbf{Id}_{VV} - \mathbf{P}_{VV}$ n'est pas inversible, on a affaire à un comportement pathologique (*i.e.* probabilité non nulle de rester indéfiniment dans les états évanescents). Dans le cas contraire, les deux expressions sont utilisables pour calculer $\mathbf{P}'$.

Ce type d'élimination des états évanescents est applicable à des modèles plus généraux (*e.g.* les réseaux de Petri stochastiques déterministes) sous certaines conditions.

Le modèle RdPSG est certainement celui qui a abouti au plus grand nombre de spécifications et d'analyses de systèmes [Ajm 95]. L'outil GreatSPN [CHI 95] a contribué à son essor. Remarquant que les analyses basées sur le graphe d'état ne requièrent qu'un graphe fini, de nombreuses extensions ont été introduites : les arcs inhibiteurs, les gardes sur les transitions, les débits dépendant du marquage courant (appelées *dépendances fonctionnelles*), etc. Nous mentionnerons si nécessaire les conséquences de ces extensions sur les méthodes plus élaborées.

3.4. Les réseaux de Petri stochastiques déterministes

Si les lois exponentielles conviennent pour la modélisation d'événements dont la distribution temporelle n'est pas connue, certaines opérations ont une durée comprise dans un intervalle de temps ou même assimilable à une constante. Dans ce cas, le choix d'une loi exponentielle conduit à des mesures très approximatives. Aussi une extension du modèle RdPSG comprenant des transitions déterministes a été introduite [Ajm 87, LIN 98]. Ces réseaux sont communément appelés réseaux de Petri stochastiques déterministes (*Deterministic Stochastic Petri Net*, DSPN). De nombreuses variantes ont été successivement proposées pour couvrir des situations différentes. Nous nous limiterons à une version de

base pour expliquer plus facilement le processus stochastique. Nous excluons les transitions immédiates. Leur traitement s'effectue par la technique de la section précédente. De même nous interdisons les dépendances fonctionnelles : en particulier les taux des transitions exponentielles et les durées des transitions déterministes sont indépendants du marquage. Le réseau fonctionne en mode *single-server* avec une politique de mémoire *enabling-memory PRD*. L'hypothèse fondamentale est qu'à tout instant *au plus une transition déterministe est franchissable*, ce qui restreint le champ d'application de ce modèle. Des travaux récents lèvent cette hypothèse au prix d'une complexité accrue.

Plusieurs méthodes ont été proposées pour l'analyse de ce type de réseaux [GER 99, LIN 98, LIN 99]. Nous nous restreignons à une des méthodes dont l'efficacité est avérée. Examinons le processus stochastique engendré par le réseau. Nous pouvons y caractériser des points de régénération :

- Chaque fois que le processus atteint un marquage où aucune transition déterministe n'est franchissable, le futur de ce processus ne dépend que du marquage.
- Chaque fois que le processus franchit une transition déterministe et qu'une transition déterministe t_k est à nouveau franchissable, le futur de ce processus ne dépend que du marquage car le délai de franchissement t_k est le délai nominal que nous notons d_k .
- Chaque fois que le processus franchit une transition exponentielle qui rend franchissable une transition déterministe t_k , le futur de ce processus ne dépend que du marquage car le délai de franchissement de la transition déterministe est d_k .

Attention dans les deux derniers cas, le point de régénération est caractérisé à la fois par le marquage atteint et par la nature du franchissement de la transition car le marquage peut être atteint par un franchissement qui ne conduit pas à un point de régénération. Pour distinguer les points de régénération des marquages, nous notons m^r , le point de régénération associé au marquage m et m^c un état atteint ayant pour marquage m et n'étant pas un point de régénération.

Il nous reste à calculer les paramètres du comportement du processus entre deux points de régénération pour appliquer les résultats de la section 2.5. Remarquons que lorsqu'on entre dans un point de régénération, on franchit une séquence de transitions exponentielles terminée éventuellement par un franchissement de la transition déterministe active. Pour un point de régénération du premier type, tout franchissement conduit à un nouveau point. Les paramètres du comportement sont donc donnés par les taux des différentes transitions franchissables.

Pour un point de régénération du deuxième ou du troisième type m^r , le franchissement des transitions exponentielles conduit à l'évolution d'une chaîne de Markov dont les états sont des m_i^c et qui se termine par :

- le franchissement de la transition déterministe d_k instants plus tard,

- un franchissement de transition exponentielle qui conduit à un m_j^r .

Appelons C_{m^r} la chaîne de Markov constituée de m^c (considéré comme l'état initial) et de ces m_i^c et m_j^r (ces derniers étant sans successeur dans la chaîne). Notons $\mathbf{Q}_{m^r}$ son générateur infinitésimal qui est directement obtenu par les taux des transitions exponentielles. On parle ici de *chaîne subordonnée*. $\pi_t^{m^r}$ est la distribution à l'instant t de cette chaîne sachant que la distribution initiale $\pi_0^{m^r}$ est concentrée en m^c . Rappelons que $\pi_t^{m^r} = \pi_0^{m^r} \cdot e^{t \cdot \mathbf{Q}_{m^r}}$.

Les probabilités de branchement entre points de régénération (matrice $\mathbf{P}$) sont déduites de l'état de cette chaîne à l'instant d_k :

$$\mathbf{P}[m^r, m_1^r] = \pi_{d_k}^{m^r}[m_1^r] + \sum_{m_2[t_k > m_1]} \pi_{d_k}^{m^r}[m_2^c]$$

qui signifie que m_1^r a été le premier point de régénération atteint avant d_k ou qu'aucun point de régénération n'a été atteint avant d_k et que le franchissement de la transition déterministe a conduit à m_1 . La durée moyenne de séjour dans un marquage avant le nouveau point de régénération est de même donnée par :

$$\text{séjour}_{m^r}(m_1) = \int_0^{d_k} \pi_t^{m^r}[m_1^c] dt$$

Pour réaliser efficacement ces calculs, différentes techniques sont possibles [JEN 53, GRO 84] mais la résolution reste cependant plus coûteuse que pour les RdPSG.

3.5. Les réseaux de Petri stochastiques à loi phase-type

Une distribution phase-type [NEU 81] est définie par une chaîne de Markov comportant un état d'absorption (*i.e.* sans successeur) et une distribution initiale. Si F en est la distribution, alors $F(t)$ est probabilité d'être dans l'état d'absorption à l'instant t . D'après la section 2.3.2, F est une distribution de probabilité si et seulement si l'état d'absorption est la seule c.f.c. puits du graphe associé à la chaîne. Dans ce cas F est définie par l'équation [5]. Les états de la chaîne (sauf le dernier) sont appelés les étages de la loi.

Il est établi qu'en un certain sens, toute distribution est une limite de distributions phase-type [COX 55]. Par exemple, une distribution exponentielle est une distribution phase-type à un étage et une distribution immédiate est une distribution phase-type sans étage. Une distribution déterministe de durée d est approchée par une distribution de n étages consécutifs de taux $\frac{n}{d}$.

Les réseaux de Petri à loi phase-type Les (RdPS-PH) ont donc un grand pouvoir d'expression stochastique. Toutefois, ils engendrent un processus stochastique de même nature que les RdPSG. En effet, le tirage d'une loi phase-type peut être vu comme un tirage probabiliste du choix du premier étage, un tirage de la loi exponentielle de l'étage, un nouveau tirage probabiliste du choix de l'étage suivant et ainsi de suite jusqu'à atteindre l'état d'absorption. Aussi, plutôt que de considérer les franchissements de transitions comme les événements du SED, on adopte un pas plus élémentaire : le changement d'étage dans la distribution. Ceci nécessite de compléter l'état du SED. Un état est défini par :

- un marquage
- pour chaque transition, un descripteur qui contient la suite ordonnée des tirages non encore retenus de la transition. Pour chacun de ces tirages, il suffit de conserver l'étage dans lequel il se trouve. Si la transition fonctionne en mode *enabling memory*, le nombre de tirages est exactement le nombre de services en cours de la transition. Si elle fonctionne en mode *age memory*, ce nombre de tirages peut être supérieur car il faut prendre en compte les services suspendus.

Chaque transition d'état qui atteint un étage d'absorption est une transition *externe* car elle modifie le marquage. Le nouveau descripteur est calculé en fonction de la politique de service et de mémoire du réseau. Les transitions *internes* ne modifient pas le marquage et dans le descripteur seul un tirage est modifié.

On peut construire le processus semi-markovien comme un graphe d'accessibilité en partant de l'état initial et en *franchissant* les transitions internes et externes. D'autres constructions plus élaborées sont possibles en remarquant que, par exemple, certains marquages sont le siège des mêmes transitions d'état.

Cependant le véritable problème est le nombre d'états du processus qui est du même ordre de grandeur que le produit de la taille de l'espace d'accessibilité et du cardinal des descripteurs. Nous verrons au chapitre 9 comment obtenir les probabilités stationnaires du réseau sans construire le processus.

4. Quelques méthodes d'analyse standard

4.1. Recherche de forme produit

Nous allons décrire cette méthode dans le cadre des RdPS à travers un exemple pour en illustrer les principes sans rentrer dans les détails algorithmiques.

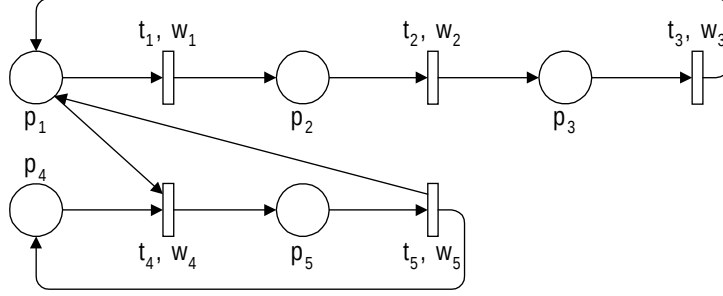


Figure 3: Un réseau de Petri à forme produit

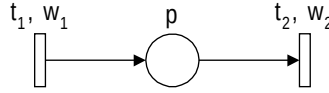


Figure 2: Modélisation d'une file d'attente par un réseau de Petri

Soit le réseau élémentaire de la figure 2, qui modélise une file d'attente. La distribution stationnaire de ce réseau (non borné) est donnée par (pour $w_1 < w_2$) :

$$\pi[n.p] = \left(1 - \frac{w_1}{w_2}\right) \cdot \left(\frac{w_1}{w_2}\right)^n$$

Dans cette équation, nous remarquons que le marquage n de la place apparaît en exposant dans la distribution stationnaire.

Aussi la première idée est de généraliser cette formule sous forme d'un produit dont les termes sont des expressions obtenues à partir des taux des k' transitions et les exposants sont les marquages des k places.

$$\pi\left[\sum_{i=1}^k n_i.p_i\right] = \frac{1}{G} \cdot \prod_{i=1}^k (f_i(w_1, \dots, w_{k'}))^{n_i}$$

G , la constante de normalisation, est définie comme la somme sur l'ensemble des marquages accessibles, des produits apparaissant dans le membre droit de l'équation.

Examinons le réseau de la figure 3. On remarque que les transitions sont partitionnables en deux sous-ensembles $T_a = \{t_1, t_2, t_3\}$ et $T_b = \{t_4, t_5\}$. A l'intérieur de chaque sous-ensemble, les préconditions de chaque transition sont les postconditions d'une autre et vice versa. Autrement dit, d'un marquage m

on peut franchir t_1 , si et seulement si m est le résultat d'un franchissement de t_3 à partir d'un autre marquage. Notons $\mathbf{Q}^a$ (resp. $\mathbf{Q}^b$) la matrice $\mathbf{Q}$ où tous les taux sont annulés exceptés ceux de T_a (resp. T_b), $\mathbf{Q} = \mathbf{Q}^a + \mathbf{Q}^b$.

La deuxième idée est de remplacer l'équation $\mathbf{X} \cdot \mathbf{Q} = 0$ par deux équations $\mathbf{X} \cdot \mathbf{Q}^a = 0$ et $\mathbf{X} \cdot \mathbf{Q}^b = 0$. Résoudre les deux systèmes n'est pas équivalent mais une solution du deuxième système nous fournira une solution du premier système. Dans ce contexte, le premier système est qualifié d'équations de balance globale et le deuxième d'équations de balance locale.

Si nous combinons ces deux idées, nous allons rechercher une forme produit comme suit :

$$\boldsymbol{\pi}(m) = \frac{1}{G} \cdot f_1(m) \cdot f_2(m)$$

où f_1 (qui dépend de w_1, w_2, w_3) n'est pas modifiée par un franchissement dans T_b et f_2 (qui dépend de w_4, w_5) n'est pas modifiée par un franchissement dans T_a . Supposons que cette forme existe et détaillons une équation de balance locale

$$\sum_{m'[t' > m, t' \in T_a]} \frac{1}{G} \cdot f_1(m') \cdot f_2(m') \cdot \mathbf{w}[t'] = \sum_{m[t'' > m'', t'' \in T_a]} \frac{1}{G} \cdot f_1(m) \cdot f_2(m) \cdot \mathbf{w}[t'']$$

Puisque $f_2(m) = f_2(m')$, l'équation est équivalente à :

$$\sum_{m'[t' > m, t' \in T_a]} f_1(m') \cdot \mathbf{w}[t'] = \sum_{m[t'' > m'', t'' \in T_a]} f_1(m) \cdot \mathbf{w}[t''] \quad [9]$$

La difficulté majeure pour parvenir à une solution est la contrainte que f_1 ne doit pas varier par un franchissement de T_b . Dans notre exemple, les marquages de p_2 et p_3 restent inchangés suite à un franchissement dans T_b et ceux de p_4 et p_5 restent inchangés suite à un franchissement dans T_a .

Nous sommes tentés d'écrire :

$$f_1(m) = (w_1)^{a_1 \cdot m(p_2) + b_1 \cdot m(p_3)} \cdot (w_2)^{a_2 \cdot m(p_2) + b_2 \cdot m(p_3)} \cdot (w_3)^{a_3 \cdot m(p_2) + b_3 \cdot m(p_3)}$$

En se rappelant que de m :

- on franchit t_1 si et seulement si m est atteint par franchissement de t_3
- on franchit t_2 si et seulement si m est atteint par franchissement de t_1
- on franchit t_3 si et seulement si m est atteint par franchissement de t_2

On essaie d'égaliser ces termes dans l'équation [9], ce qui donne après simplification :

$$\begin{aligned} (w_1)^{b_1} \cdot (w_2)^{b_2} \cdot (w_3)^{b_3+1} &= w_1 \\ (w_1)^{-a_1+1} \cdot (w_2)^{-a_2} \cdot (w_3)^{-a_3} &= w_2 \\ (w_1)^{a_1-b_1} \cdot (w_2)^{a_2-b_2+1} \cdot (w_3)^{a_3-b_3} &= w_3 \end{aligned}$$

La seule solution (pour des taux quelconques) est alors :

$$f_1(m) = \left(\frac{w_1}{w_2}\right)^{m(p_2)} \cdot \left(\frac{w_1}{w_3}\right)^{m(p_3)}$$

De manière analogue :

$$f_2(m) = \left(\frac{w_4}{w_5}\right)^{m(p_4)}$$

Dans le cas général des conditions d'existence de cette décomposition sont satisfaites pour une sous-classe de réseaux appelés réseaux de Petri à forme produit (RdPS-FP) (*Product Form Stochastic Petri Net*, PF-SPN) [HEN 90]. De plus, [HAD 01] établit une condition nécessaire et suffisante purement structurelle pour qu'un RdPS admette une forme produit quels que soient ses paramètres stochastiques. Il reste encore un point délicat. Si le calcul de la constante de normalisation est réalisé d'une manière naïve, ceci revient à énumérer les états accessibles et donc réduire l'intérêt de la méthode. Heureusement la présence d'invariants caractérisant l'espace d'accessibilité simplifie ce calcul [SER 93].

Pour conclure, les méthodes à base de forme produit ont une complexité de calcul très faible mais elles ne peuvent s'appliquer qu'à des modèles dont le pouvoir d'expression ne couvre pas des synchronisations complexes.

4.2. Calcul de bornes

Les bornes stochastiques que nous allons exhiber sont valables pour toutes les distributions (de moyenne finie) des délais des transitions [CHI 93]. Elles reposent uniquement sur l'existence d'une distribution stationnaire des marquages et des débits des transitions. Ces bornes sont donc valables entre autres pour les réseaux présentés dans ce chapitre y compris les réseaux à espace d'états infinis. En revanche, ces bornes ne seront de bonne qualité que lorsque les mesures de performance seront insensibles (*i.e.* ne dépendent que de la moyenne des distributions) ou faiblement sensibles. Cette analyse est étendue dans [LIU 95] où l'auteur intègre des contraintes liées à la variance des distributions.

L'idée générale est :

- de représenter par des variables les indices de performance,
- d'établir des contraintes linéaires entre ces variables,
- de maximiser ou minimiser, par programmation linéaire, une fonction linéaire des variables (qui représente l'indice de performance à évaluer) soumises aux contraintes précédentes.

De nombreux algorithmes sont possibles pour cette dernière étape dont les plus efficaces s'exécutent en temps polynômial par rapport à la taille des contraintes [NEM 89]. Nous nous limiterons aux deux premiers points.

Pour chaque place p , la variable $\bar{m}_p$ dénote le marquage moyen de p . Pour chaque transition t , la variable $\bar{\chi}_t$ dénote le débit moyen de t . Enfin la variable σ_t est le nombre d'occurrences de t dans une pseudo-séquence de franchissements puisque ce nombre n'est pas nécessairement un entier. Cette dernière famille permet de définir des contraintes : elle n'interviendra pas dans la fonction à maximiser ou minimiser.

La première contrainte est naturellement définie par :

$$\forall p, \forall t, \quad \bar{m}_p \geq 0, \quad \bar{\chi}_t \geq 0, \quad \sigma_t \geq 0$$

Le marquage moyen $\bar{m}$ est une moyenne pondérée par π des marquages accessibles m , chacun d'entre eux étant atteint par une séquence σ_m (de vecteur d'occurrences $\vec{\sigma}_m$) à partir de m_0 . D'où :

$$\bar{m} = m_0 + \sum_{m \text{ accessible}} \pi_m \cdot (\mathbf{Post}[p, t] - \mathbf{Pré}[p, t]) \cdot \vec{\sigma}_m$$

Le deuxième terme est le produit de la matrice d'incidence par une moyenne pondérée de séquences que nous pouvons remplacer par nos variables σ_t . Ce qui nous donne comme contrainte :

$$\bar{m}_p = m_0(p) + \sum_t (\mathbf{Post}[p, t] - \mathbf{Pré}[p, t]) \cdot \sigma_t$$

Les autres contraintes sont de deux types : les premières sont obtenues en analysant la situation à l'équilibre et les secondes en établissant une relation sur le comportement du processus jusqu'à un instant donné et en étudiant cette relation lorsque l'instant tend vers l'infini. Cette dernière technique se nomme analyse opérationnelle.

Nous exhibons quelques exemples de chacun des deux types de contraintes. Supposons que deux transitions t et t' soient simultanément franchissables ou infranchissables et que les probabilités de choix entre t et t' restent constantes. C'est le cas par exemple lorsque les transitions sont exponentielles de taux constant et *single-server*. Appelons r_t et $r_{t'}$ ces probabilités. De manière évidente :

$$\frac{\bar{\chi}_t}{r_t} = \frac{\bar{\chi}_{t'}}{r_{t'}}$$

Puisque la distribution stationnaire du marquage existe, les flux entrants d'une place doivent être égaux aux flux sortants. D'où :

$$\forall p, \sum_t \mathbf{Pré}[p, t] \cdot \bar{\chi}_t = \sum_t \mathbf{Post}[p, t] \cdot \bar{\chi}_t$$

Soit une transition t dont le délai a une moyenne $\frac{1}{w_t}$, fonctionnant en mode *infinite-server* et qui n'a qu'un arc en entrée de valuation 1 connecté à la place p qui n'a que t pour sortie. Observons le processus entre 0 et θ . Soit $m_p(s)$ le nombre de jetons présents à l'instant s , soit $\bar{m}_p(\theta)$ le nombre moyen de jetons entre 0 et θ :

$$\bar{m}_p(\theta) = \frac{1}{\theta} \int_0^\theta m_p(s) ds$$

Notons $V(\theta)$ la somme des durées de présence des jetons en p qui sont consommés avant θ et $U(\theta)$ la somme des durées de présence des jetons en p qui sont produits avant θ ou présents dans le marquage initial.

Pour établir une relation entre ces quantités, supposons les jetons clients de la transition payant un coût uniforme de présence de taux 1 : autrement dit un jeton présent dans un intervalle ds rapporte ds francs. Si les jetons arrivés avant θ payent leur présence jusqu'à l'instant θ , $\int_0^\theta m_p(s) ds$ est la somme gagnée par la transition entre 0 et θ . Si les jetons arrivés avant θ payent leur présence à leur départ, $V(\theta)$ est la somme gagnée par la transition entre 0 et θ . Si les jetons arrivés avant θ payent leur présence à leur arrivée, $U(\theta)$ est la somme gagnée par la transition entre 0 et θ . D'où :

$$V(\theta) \leq \theta \cdot \bar{m}_p(\theta) \leq U(\theta) \Leftrightarrow \frac{V(\theta)}{\theta} \leq \bar{m}_p(\theta) \leq \frac{U(\theta)}{\theta} \quad [10]$$

Analysons le comportement à l'infini. Soit $N(\theta)$ le nombre jetons arrivés entre 0 et θ . Puisque le réseau a une distribution stationnaire, le flux entrant de jetons en p doit être égal au flux sortant.

$$\lim_{\theta \rightarrow \infty} \frac{N(\theta)}{\theta} = \bar{\chi}_t$$

Puisqu'il y a une distribution stationnaire du marquage,

$$\lim_{\theta \rightarrow \infty} \bar{m}_p(\theta) = \bar{m}_p$$

Enfin chaque jeton a un séjour égal au délai de franchissement de la transition. Appelons d_n la durée du séjour du $n^{ième}$ jeton. D'après la loi des grands nombres :

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n d_i}{n} = \frac{1}{w_t}$$

Il nous reste à établir une relation entre ces quantités.

$$\frac{\bar{\chi}_t}{w_t} = \lim_{\theta \rightarrow \infty} \left(\frac{N(\theta)}{\theta} \right) \cdot \left(\frac{\sum_{i=1}^{N(\theta)} d_i}{N(\theta)} \right) = \lim_{\theta \rightarrow \infty} \frac{U(\theta)}{\theta}$$

À l'aide d'un raisonnement analytique [STI 74] on démontre que :

$$\lim_{\theta \rightarrow \infty} \frac{U(\theta)}{\theta} = \lim_{\theta \rightarrow \infty} \frac{V(\theta)}{\theta}$$

Par passage à la limite, l'équation [10] nous permet d'obtenir une nouvelle contrainte :

$$\frac{\bar{\chi}_t}{w_t} = \bar{m}_p$$

Cette contrainte est une des variantes de la formule de Little. Notons que le calcul de bornes a aussi été appliqué aux réseaux bien formés, car les symétries de ce modèle facilitent l'expression des contraintes.

4.3. Méthodes approximatives

4.3.1. Approximation par décomposition

Nous nous intéressons ici à une sous-classe de RdPS appelés graphes marqués stochastiques (GMS) (*Stochastic Marked Graph*, SMG) pour laquelle le RdP ordinaire est un graphe marqué (*c.f.* ch. 3). Cette classe est souvent utilisée pour des modélisations de flux comme par exemple dans les ateliers flexibles [SIL 97, VAL 97].

Dans ces réseaux, chaque place est sortie d'une unique transition et entrée d'une unique transition. De plus le réseau, vu comme un graphe, est fortement connexe. Nous supposons que le réseau fonctionne en mode *infinite-server* sachant que les autres modes se traitent de façon analogue.

Tous les débits des transitions en régime stationnaire (*i.e.* le nombre de franchissements par unité de temps) sont égaux. En effet, un chemin joint chaque couple de transitions (t_1, t_2) et les (nouvelles) marques sur ce chemin sont produites par t_1 et consommées par t_2 . Si le débit de t_1 était supérieur à celui de t_2 , le nombre de marques dans ce chemin tendrait vers l'infini ce qui est impossible puisque ce nombre est borné par le nombre initial de marques d'un circuit englobant le chemin. Par symétrie, on en déduit que le débit de t_1 est égal à celui de t_2 . On parle donc de débit du réseau. Pour calculer le débit du réseau, on pourrait établir la distribution stationnaire du réseau, choisir une transition et appliquer la formule [8].

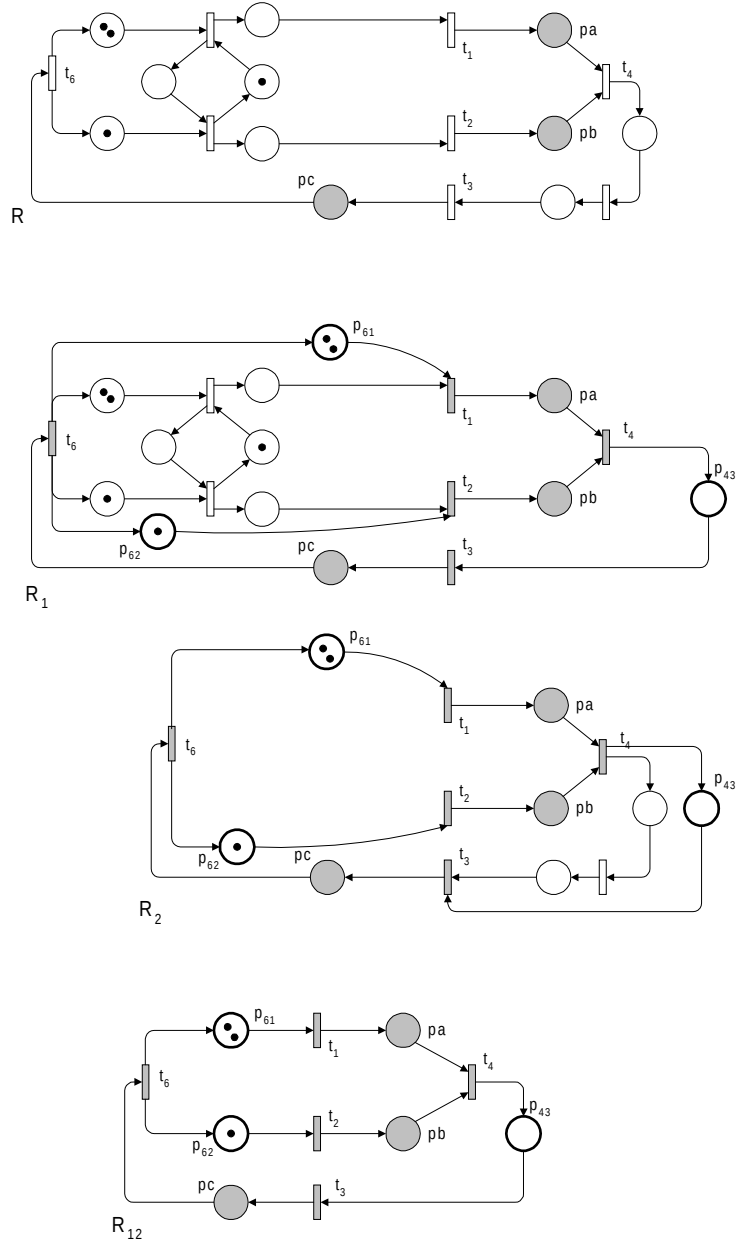


Figure 4: Décomposition et abstraction d'un GMS

L'objectif de la méthode d'approximation par décomposition est de substituer à la construction du graphe d'accessibilité de ce réseau, la construction de graphes pour des sous-réseaux obtenus par décomposition [CAM 94]. En effet, pour des décompositions bien choisies, la taille du graphe total est de l'ordre du produit des tailles des graphes des sous-réseaux. L'algorithme se déroule en deux étapes :

- la décomposition en sous-réseaux
- le calcul approché du débit

Pour obtenir une décomposition, on choisit un ensemble de places appelé une coupure qui partage le réseau en deux composantes connexes. Le choix de la coupure doit être guidé par la sémantique du système. En règle générale, on cherchera à minimiser la taille de la coupure. Sur le schéma 4, la coupure est donnée par les places pa , pb et pc . Soient R_1 et R_2 les deux sous-réseaux ; chacun d'entre eux est d'abord constitué d'une composante connexe, de la coupure et des transitions connexes à la coupure. Pour compléter les réseaux, les places p_{61} , p_{62} et p_{43} sont ajoutées aux deux réseaux. Chaque place correspond à un couple de transitions formant la frontière d'un sous-réseau. Par exemple, p_{61} correspond au couple (t_6, t_1) .

Les places p_{61} , p_{62} (resp. p_{43}) forment une abstraction de la première (resp. deuxième) composante. Le marquage initial de p_{61} est choisi comme le minimum du marquage d'un chemin de t_6 à t_1 dans cette composante. Il en va de même des autres places.

Le réseau R_{12} constitue une abstraction complète du réseau initial. En raison des propriétés structurelles des GMS, toutes ces abstractions préservent les langages (en masquant les transitions absentes de l'abstraction) et l'ensemble des marquages accessibles (en oubliant les places absentes de l'abstraction).

Le calcul du débit du réseau se fait de manière itérative. Nous l'expliquons sur l'exemple. La vitesse de t_3 dans le réseau R_1 sera ajustée à chaque tour de boucle pour refléter l'activité de la deuxième composante. Il en sera de même pour les vitesses de t_1 et t_2 dans R_2 . On choisit initialement une valeur de vitesse pour t_3 dans R_1 , par exemple sa valeur dans R . Puis chaque tour de boucle comporte quatre étapes :

- à partir de la distribution stationnaire, on calcule le débit du réseau R_1 , appelons-le χ_1 , et le marquage moyen de p_{61} et p_{62} . On note que ces marquages moyens sont proportionnels aux temps de *service* de t_1 et de t_2 pour ces places (*i.e.* le temps moyen de consommation d'un jeton de p_{61} et p_{62}) puisque la production est simultanée et que nous avons choisi une politique *infinite-server*.
- le ratio des vitesses de t_1 et t_2 dans R_2 est donc fixé par l'étape précédente. Il reste à calculer le facteur d'échelle. Ceci se fait dans R_{12} où pour des valeurs successives de ce facteur d'échelle, on calcule la distribution stationnaire et le débit du réseau pour se rapprocher le plus possible de

- χ_1 . Le graphe de R_{12} est de taille très réduite par rapport aux autres graphes. Aussi cette étape reste d'une complexité satisfaisante.
- cette étape est symétrique de la première et l'on calcule le débit χ_2 de R_2 avec les vitesses de t_1 et t_2 obtenues par l'étape précédente. Dans cet exemple, puisque t_3 est la seule transition qui conduit de R_2 à R_1 , le calcul des proportions est inutile.
 - on retourne au réseau R_{12} pour calculer la vitesse de t_3 à utiliser dans R_1 . On essaie des valeurs successives pour que le débit de R_{12} se rapproche le plus possible de χ_2 .

On sort de la boucle lorsque les valeurs χ_1 et χ_2 sont suffisamment proches pour supposer qu'elles correspondent au débit du réseau R . Il n'y a pas de garantie théorique de convergence, ni de précision du résultat obtenu. Ceci est d'ailleurs vrai pour presque toutes les méthodes approximatives. Cependant, les expérimentations montrent une convergence très rapide (moins de 10 tours de boucle) et une erreur inférieure à 1%. Ces bons résultats sont certainement dûs au fait que l'approximation quantitative est basée sur une décomposition fonctionnelle valide.

4.3.2. Approximation par valeurs moyennes

Cette analyse s'applique a priori sur des réseaux de Petri stochastiques quelconques bien que les outils développés reconnaissent les distributions uniformes, exponentielles, déterministes et certaines de leurs combinaisons. Nous la présentons lorsque les transitions fonctionnent en mode *single-server* et *enabled memory PRD*.

La méthode s'articule autour de la construction d'un graphe dit *graphe d'états probabilisé* [JUA 91]. Chaque noeud de ce graphe est composé d'un marquage et de distributions associées à chaque transition. Le premier noeud est composé du marquage initial et des distributions nominales.

On détermine les transitions franchissables $\{t_k\}_{k \in K}$ et on calcule la probabilité de franchissement de chaque transition. Nous supposerons dans la suite que les distributions sont continues. Il s'agit alors de déterminer la probabilité que le tirage de chaque transition soit le plus petit et ceci se calcule par :

$$\Pr(t_k \text{ franchie}) = \int_0^\infty \prod_{k' \neq k} (1 - F_{k'}(s)) F_k\{ds\}$$

où les F_k sont les distributions associées aux t_k . Dans le cas général de distributions non continues, il faut intégrer les paramètres de post-sélection ce qui complique les expressions mais ne modifie pas la nature du problème.

La durée moyenne du séjour dans le noeud est donnée de manière analogue par :

$$\int_0^\infty \prod_{k \in K} (1 - F_k(s)) ds$$

On construit ensuite un successeur par franchissement possible de transition. Ici se situe l'approximation car on considère que la transition est franchie de manière déterministe avec un temps moyen qui se calcule par :

$$\theta_k = \frac{1}{\Pr(t_k \text{ franchie})} \int_0^\infty s \cdot \prod_{k' \neq k} (1 - F_{k'}(s)) F_k\{ds\}$$

Il y a ici identification d'une variable aléatoire à sa moyenne. Les nouvelles distributions des transitions encore franchissables sont alors :

$$F'_{k'}(t) = \frac{F_{k'}(t + \theta_k) - F_{k'}(\theta_k)}{1 - F_{k'}(\theta_k)}$$

Les autres distributions sont de nouveau les distributions nominales. Si le graphe d'accessibilité et les distributions intermédiaires sont finies, alors le graphe probabilisé est fini. On peut établir des conditions suffisantes sur le réseau de Petri qui garantissent cette propriété. Dans le cas où le graphe est infini, un mécanisme d'arrêt de développement de branches est introduit qui tient compte des probabilités de branchement pour éliminer des noeuds supposés de faible probabilité.

Pour tirer profit de ce graphe, on l'interprète comme un processus régénératif donné par les probabilités et les durées moyennes de séjour calculées plus haut. La résolution s'effectue alors comme indiqué en section 2.5.1.

Il est malheureusement difficile de cerner les critères qui garantiront une bonne approximation. On peut penser que de nombreuses distributions unimodales diminueront la qualité de l'approximation. D'autre part, certains états accessibles peuvent être absents du graphe probabilisé (même sans le mécanisme de coupure des branches).

4.4. Réseaux de Petri non bornés

Jusqu'à présent, la plupart des méthodes que nous avons présentées s'appliquaient à des réseaux bornés. Nous terminons ce chapitre par une méthode exacte qui s'applique aux systèmes infinis [FLO 89, HAV 95]. Nous nous restreignons ici aux RdPS mais cette analyse reste valable moyennant des adaptations aux RdPSG et même aux RdPS-PH. Lorsque les réseaux ont une unique place

p non bornée, il est possible de calculer la distribution stationnaire en imposant de faibles restrictions sur le réseau (d'autres restrictions sont possibles) :

- Les arcs connectés à la place p ont une valuation 1.
- Deux valeurs quelconques de $m(p)$ supérieures à un seuil k_0 induisent les mêmes débits des transitions et les mêmes conditions de franchissement données dans les dépendances fonctionnelles.

Dans ce cas, les marquages accessibles sont partitionnés selon le marquage de p en une famille $\{S_k\}_{k=0}^{\infty}$ où S_k est le sous-ensemble des marquages accessibles tels que le marquage de p soit k . Nos contraintes impliquent que de S_k , soit on atteint S_{k-1} ou S_{k+1} soit on reste dans S_k . De même, si le marquage de p est supérieur à k_0 , la place p n'a plus d'influence sur le comportement du réseau. Aussi la matrice $\mathbf{Q}$ du générateur infinitésimal du réseau présente au delà de k_0 des régularités qui se traduisent par l'existence de trois matrices :

- $\mathbf{A}_0$ est la sous-matrice de transition de S_k vers S_{k+1}
- $\mathbf{A}_1$ est la sous-matrice de transition de S_k vers S_k
- $\mathbf{A}_2$ est la sous-matrice de transition de S_k vers S_{k-1}

Supposons que la chaîne soit irréductible et ergodique avec une distribution stationnaire π et notons π_k la distribution sur les états S_k . L'équation d'équilibre se réécrit pour $k > k_0$:

$$\pi_k \cdot \mathbf{A}_0 + \pi_{k+1} \cdot \mathbf{A}_1 + \pi_{k+2} \cdot \mathbf{A}_2 = 0$$

Nous voulons établir une relation de récurrence entre π_k et π_{k+1} . D'après la structure de $\mathbf{Q}$, $\mathbf{A}_1$ est inversible et $-\mathbf{A}_1^{-1}$ est une matrice positive. D'où :

$$\pi_{k+1} + \pi_k \cdot \mathbf{A}_0 \cdot \mathbf{A}_1^{-1} = -\pi_{k+2} \cdot \mathbf{A}_2 \mathbf{A}_1^{-1} \geq 0$$

On améliore la relation en définissant une suite de matrices croissantes :

$$\mathbf{R}_0 = -\mathbf{A}_0 \cdot \mathbf{A}_1^{-1} \text{ et } \mathbf{R}_{n+1} = -(\mathbf{A}_0 + (\mathbf{R}_n)^2 \cdot \mathbf{A}_2) \cdot \mathbf{A}_1^{-1}$$

On démontre par récurrence que le terme gauche ci-dessous reste toujours positif (et décroît puisque les $\mathbf{R}_n$ sont croissants). En effet :

$$\begin{aligned} \pi_{k+1} - \pi_k \cdot \mathbf{R}_{n+1} &= \pi_{k+1} + \pi_k \cdot \mathbf{A}_0 \cdot \mathbf{A}_1^{-1} + \pi_k \cdot (\mathbf{R}_n)^2 \cdot \mathbf{A}_2 \cdot \mathbf{A}_1^{-1} \\ &= -(\pi_{k+2} - \pi_k \cdot (\mathbf{R}_n)^2) \cdot \mathbf{A}_2 \mathbf{A}_1^{-1} \\ &= -(\pi_{k+2} - \pi_{k+1} \cdot \mathbf{R}_n + (\pi_{k+1} - \pi_k \cdot \mathbf{R}_n) \cdot \mathbf{R}_n) \cdot \mathbf{A}_2 \mathbf{A}_1^{-1} \geq 0 \end{aligned}$$

Puisque toutes les composantes de π_k sont non nulles, les matrices $\mathbf{R}_n$ sont bornées et la suite converge vers une matrice $\mathbf{R}$ qui vérifie (en passant à la limite) :

$$\pi_{k+1} - \pi_k \cdot \mathbf{R} \geq 0 \text{ pour } k > k_0 \quad [11]$$

$$\mathbf{R} = -(\mathbf{A}_0 + \mathbf{R}^2 \cdot \mathbf{A}_2) \cdot \mathbf{A}_1^{-1} \quad [12]$$

Définissons un vecteur $\boldsymbol{\pi}'$ par :

$$\forall k \leq k_0, \quad \boldsymbol{\pi}'_k = \boldsymbol{\pi}_k \quad \text{et} \quad \forall k > k_0, \quad \boldsymbol{\pi}'_k = \boldsymbol{\pi}_{k_0+1} \cdot \mathbf{R}^{(k-k_0-1)}$$

Alors, d'après l'équation [12], ce vecteur est solution de l'équation $\mathbf{X} \cdot \mathbf{Q} = 0$ et d'après l'équation [11], il est inférieur ou égal à $\boldsymbol{\pi}$ composante par composante. Donc la somme de ses composantes est finie. En le normalisant (*i.e.* en le divisant par cette somme), on obtient $\boldsymbol{\pi}''$ une distribution solution. Or la distribution solution est unique. Donc $\boldsymbol{\pi} = \boldsymbol{\pi}'' = \boldsymbol{\pi}'$.

L'équation de normalisation s'écrit alors :

$$\sum_{k=0}^{k_0} \boldsymbol{\pi}_k \cdot \mathbf{1}^T + \sum_{k=k_0+1}^{\infty} \boldsymbol{\pi}_{k_0+1} \cdot \mathbf{R}^{(k-k_0-1)} \cdot \mathbf{1}^T = \sum_{k=0}^{k_0} \boldsymbol{\pi}_k \cdot \mathbf{1}^T + \boldsymbol{\pi}_{k_0+1} \cdot (\mathbf{Id} - \mathbf{R})^{-1} \cdot \mathbf{1}^T = 1$$

Une fois calculées les matrices $\mathbf{R}$ (*e.g.* en l'approchant par les $\mathbf{R}_n$) et $(\mathbf{Id} - \mathbf{R})^{-1}$, nous sommes en mesure de nous ramener à une résolution linéaire dans un espace fini. Le système en question est $\mathbf{X} \cdot \mathbf{Q} = 0$, réduit aux états de $\{S_k\}_{k \leq k_0+1}$ et complété par l'équation de normalisation [HAV 98].

Ce procédé a aussi été employé avec succès pour approximer des systèmes finis dans lesquels une place atteint de grandes valeurs. Le seuil k_0 est souvent très inférieur à la borne de la place. Si de plus le système vérifie un ensemble de conditions de quasi-réversibilité [KEL 79], cette approximation devient un résultat exact [HAV 93].

5. Conclusion

Les réseaux de Petri stochastiques ont été initialement introduits comme un autre moyen de représenter les systèmes à événements discrets stochastiques à lois exponentielles. Au cours des vingt dernières années, les besoins en modélisation ont amené à étendre ce modèle à des lois plus générales (déterministes, phase type, quelconque) tout en incluant les possibilités de choix à durée nulle. Ces extensions engendrent des familles de réseau de Petri stochastique dont les propriétés dépendent de plusieurs choix, parfois assez fins, liés à la sémantique stochastique du modèle. Il reste que, pour la majeure partie d'entre-eux, les processus stochastiques engendrés par ces réseaux sont dérivés des processus de renouvellement. On a également développé des méthodes appropriées d'analyse de ces processus. Certaines tentent de tirer partie des résultats obtenus dans le cadre de la théorie des réseaux de files d'attente. Pour la plupart cependant,

elles s'appuient, dans une large mesure, sur des propriétés que l'on peut établir pour les réseaux de Petri ordinaires. Les deux chapitres suivant présentent deux exemples d'approches caractéristiques de cette démarche : les réseaux de Petri stochastiques bien formés, extension stochastique des réseaux colorés bien formés (chapitre 10), et les méthodes d'analyse tensorielle des réseaux de Petri stochastiques (chapitre 11).

Références

- [Ajm 84] M. AJMONE MARSAN, G. BALBO ET G. CONTE. A class of generalized of stochastic Petri nets for the performance evaluation of multiprocessor systems. *ACM Transactions on computer systems*, 2(2) :93–122, May 1984.
- [Ajm 87] M. AJMONE MARSAN ET G. CHIOLA. On Petri nets with deterministic and exponentially distributed firing times. In G. ROZENBERG, ed, *Advances in Petri Nets 1987*, number 266 in LNCS, pages 132–145. Springer-Verlag, 1987.
- [Ajm 95] M. AJMONE MARSAN, G. BALBO, G. CONTE, S. DONATELLI ET G. FRANCESCHINIS. *Modelling with Generalized Stochastic Petri Nets*. Wiley series in parallel computing. John Wiley & Sons, England, 1995.
- [BOL 98] G. BOLCH, S. GREINER, H. DE MEER ET K.S. TRIVEDI. *Queueing networks and Markov chains*. John Wiley & Sons, New-York, 1998.
- [CAM 94] J. CAMPOS, J.M. COLOM, H. JUNGNITZ ET M. SILVA. Approximate throughput computation of stochastic marked graphs. *IEEE Transactions on Software Engineering*, 20(7) :526–535, July 1994.
- [CHI 93] G. CHIOLA, C. ANGLANO, J. CAMPOS, J.M. COLOM ET M. SILVA. Operational analysis of timed Petri nets and application to computation of performance bounds. In *Proc. of the 5th International Workshop on Petri Nets and Performance Models*, pages 128–137, Toulouse, France, October 19–22 1993. IEEE Computer Society Press.
- [CHI 95] G. CHIOLA, G. FRANCESCHINIS, R. GAETA ET M. RIBAUDO. GreatSPN 1.7 : GGraphical Editor and Analyser for Timed and Stochastic Petri nets. *Performance Evaluation*, 24(1,2) :47–68, 1995.
- [COX 55] D. R. COX. A use of complex probabilities in the theory of stochastic processes. *Proc. Cambridge Philosophical Society*, pages 313–319, 1955.
- [DAL 97] Y. DALLERY, Y. LIU ET D. TOWSLEY. Properties of fork/join queueing networks with blocking under various operating mechanisms. *IEEE Transactions on Robotics and Automation*, 13(4) :503–518, 1997.
- [FDI 86] S. FDIDA, G. PUJOLLE ET D. MAILLES. Réseaux de files d’attente avec sémaphores. 5(3) :187–196, mai-juin 1986.
- [FEL 68] W. FELLER. *An introduction to probability theory and its applications. Volume I*. John Wiley & Sons, 1968. (third edition).
- [FEL 71] W. FELLER. *An introduction to probability theory and its applications. Volume II*. John Wiley & Sons, 1971. (second edition).
- [FLO 78] G. FLORIN ET S. NATKIN. Évaluation des performances d’un protocole de communication à l’aide des réseaux de Petri et des processus stochastiques. In *Journées AFCET Multi-ordinateurs, multi-processeurs en temps réel*, CNRS, Paris, France, mai 1978.
- [FLO 85] G. FLORIN ET S. NATKIN. Les réseaux de Petri stochastiques. 4(1) :143–160, 1985.
- [FLO 89] G. FLORIN ET S. NATKIN. Necessary and sufficient ergodicity condition for open synchronized queueing networks. *IEEE Transactions on software engineering*, 15(4) :367–380, 1989.
- [GER 99] R. GERMAN ET M. TELEK. Formal relation of Markov renewal theory and supplementary variables in the analysis of stochastic Petri nets. In P. BUCHHOLZ ET M. SILVA, eds, *Proc. of the 8th International Workshop on Petri Nets and Performance Models*, pages 64–73, Zaragoza, Spain, September 8–10 1999. IEEE Computer Society Press.

- [GRO 84] D. GROSS ET D.R. MILLER. The randomization technique as a modeling tool an solution procedure for transient markov processes. *Oper. Res.*, 32 :345–361, 1984.
- [HAD 01] S. HADDAD, P. MOREAUX, M. SERENO ET M. SILVA. Structural characterization and behavioural properties of product form stochastic Petri nets. In *Proc. of the 22th International Conference on Application and Theory of Petri Nets*, NewCastle Upon Tine, UK., June 25–29 2001. To appear.
- [HAV 93] B.R. HAVERKORT. Approximate performability and dependability modeling using generalized stochastic Petri nets. *Performance Evaluation*, 18(1) :61–78, 1993.
- [HAV 95] B.R. HAVERKORT. Matrix-geometric solution of infinite stochastic Petri nets. In *Proc. of the International Performance and Dependability Symposium*, pages 72–81. IEEE Computer Society Press, 1995.
- [HAV 98] B.R. HAVERKORT. *Performance of Computer Communication Systems*. John Wiley & Sons, 1998.
- [HEN 90] W. HENDERSON, C.E.M. PEARCE, P.G. TAYLOR ET N.M. VAN DIJK. Closed queueing networks with batch services. *Queueing Systems*, (6) :59–70, 1990.
- [HIL 96] J. HILLSTON. *A Compositional Approach to Performance Modelling*. Cambridge University Press, 1996.
- [JEN 53] A. JENSEN. Markov chains as an aid in the study of Markov processes. *Skand. Aktuarietidskrift*, 3 :87–91, 1953.
- [JUA 91] G. JUANOLE ET Y. ATAMNA. Dealing with arbitrary time distributions with the stochastic timed Petr net model – application to queueing systems. In *Proc. of the fourth International Workshop on Petri Nets and Performance Models*, pages 32–51, Melbourne, Australia, December 2–5 1991. IEEE Computer Society Press.
- [KEL 79] F.P. KELLY. *Reversibility and stochastic networks*. John Wiley & Sons, England, 1979.
- [KLE 75] L. KLEINROCK. *Queueing systems. Volume I : Theory*. Wiley-Interscience, New-York, 1975.
- [LAP 95] J.C. LAPRIE, ed. *Guide de la sûreté de fonctionnement*. Cépaduès - Éditions, Toulouse, France, 1995.
- [LIN 98] C. LINDEMANN. *Performance Modelling with Deterministic and Stochastic Petri Nets*. John Wiley & Sons, 1998.
- [LIN 99] C. LINDEMANN, A. REUYS ET A. THÜMLER. DSPNexpress 2.000 performance and dependability modeling environment. In *Proc. of the 29th Int. Symp. on Fault Tolerant Computing*, Madison, Wisconsin, June 1999.
- [LIU 95] Z. LIU. Performance bounds for stochastic timed Petri nets. In *Proc. of the 16th International Conference on Application and Theory of Petri Nets*, number 935 in LNCS, pages 316–334, Turin, Italy, June 26–30 1995. Springer-Verlag.
- [MEY 80] J.F. MEYER. On evaluating the performability of degradable computing systems. *IEEE Transactions on Computers*, 29(8) :720–731, August 1980.
- [MOL 81] M. K. MOLLOY. *On the integration of delay and throughput in distributed processing models*. PhD dissertation, University of California, Los Angeles, CA, USA, September 1981.

- [NEM 89] G.L. NEMHAUSER, A.H.G. RINNOOY KAN ET M.J. TODD, eds. *Optimization*, volume 1 of *Handbooks in Operations research and Management Science*. North-Holland, Amsterdam, 1989.
- [NEU 81] M. F. NEUTS. *Matrix-geometric solutions in stochastic models - an algorithmic approach*. The John Hopkins University Press, London, 1981.
- [REI 98a] W. REISIG ET G. ROZENBERG, eds. *Lectures on Petri Nets I : Basic models*. Number 1491 in LNCS. Springer-Verlag, June 1998. Advances in Petri nets.
- [REI 98b] W. REISIG ET G. ROZENBERG, eds. *Lectures on Petri Nets II : Applications*. Number 1492 in LNCS. Springer-Verlag, June 1998. Advances in Petri nets.
- [SER 93] M. SERENO ET G. BALBO. Computational algorithms for product form solution stochastic Petri nets. In *Proc. of the 5th International Workshop on Petri Nets and Performance Models*, pages 98–107, Toulouse, France, October 19–22 1993. IEEE Computer Society Press.
- [SIL 97] M. SILVA ET E. TERUEL. Petri nets for the design and operation of manufacturing systems. *European journal of control*, 3(3), 1997.
- [STE 94] W. J. STEWART. *Introduction to the numerical solution of Markov chains*. Princeton University Press, USA, 1994.
- [STI 74] S. STIDHAM. A last word on $L = \lambda W$. *Operations research*, 22(2) :417–421, 1974.
- [TRI 82] K. S. TRIVEDI. *Probability & statistics with reliability, queueing, and computer science applications*. Prentice Hall, Englewood Cliffs, NJ, USA, 1982.
- [TRI 92] K. S. TRIVEDI, J. K. MUPPALA, S. P. WOOLE ET B. R. HAVERKORT. Composite performance and dependability analysis. *Performance Evaluation*, 14(3–4) :197–215, February 1992.
- [VAL 97] R. VALETTE. Some issues about Petri net application to manufacturing and process supervisory control. In *Proc. of the 18th International Conference on Application and Theory of Petri Nets*, number 1248 in LNCS, pages 23–41, Toulouse, France, June 23–27 1997. Springer-Verlag.

Index

Haddad, S., 1
Moreaux, P., 1
analyse opérationnelle, 25
chaîne incluse, 10
distribution stationnaire, 5
graphe d'états probabilisé, 30

bornes stochastiques, 24

chaîne irréductible, 8
chaîne de Markov à temps continu,
9
chaîne de Markov à temps discret,
7

forme produit, 22

marquages tangibles, 17
marquages évanescents, 17
méthode d'approximation par dé-
composition, 27

processus *ergodique*, 5
processus de Markov régénératif, 12
processus de renouvellement, 5
processus semi-markovien, 11

réseau de Petri stochastique géné-
ralisé, 17
réseaux de Petri stochastiques dé-
terministes, 18
réseaux de Petri à loi phase-type,
20