Échantillonnage préférentiel pour le model checking statistique

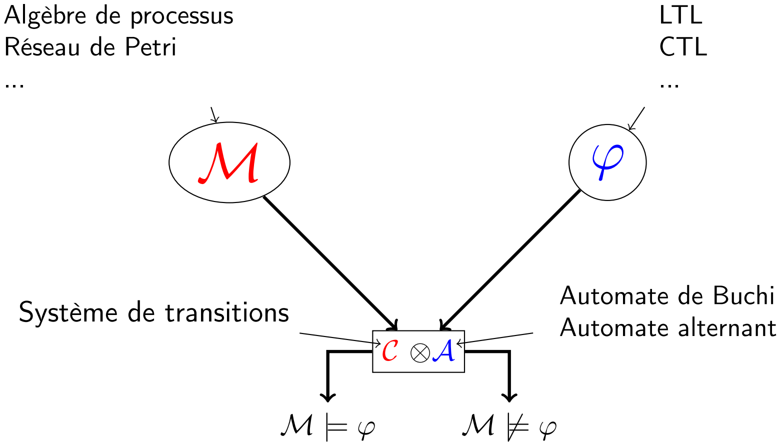
Benoît Barbot, Serge Haddad et Claudine Picaronny

LSV, CNRS & ENS de Cachan

Vendredi 4 novembre 2011

Plan

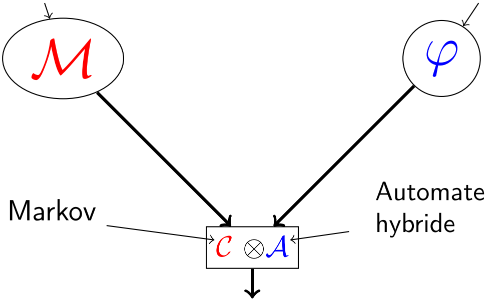
Model checking



Model checking de systèmes probabilistes

Algèbre de processus stochastique
Réseau de Petri stochastique
...

PCTL
HASL
...



Chaîne de Markov

Automate *déterministe* hybride

$$p = \Pr(\mathcal{M} \models \varphi)$$

Approches numérique et statistique

- Approche numérique
  - Logiques arborescentes (basées sur CTL)
  - Valeur exacte (aux approximations numériques près)
  - Implémenté efficacement dans de nombreux outils (PRISM, MRMC, GreatSPN)
  - Hypothèses probabilistes fortes
  - Taille de l'espace mémoire proportionnelle à la taille du processus stochastique

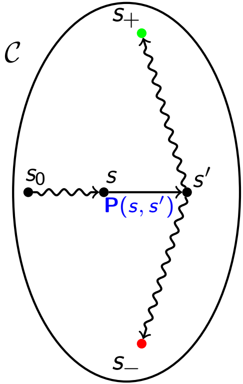
Approches numérique et statistique

- Approche numérique
  - Logiques arborescentes (basées sur CTL)
  - Valeur exacte (aux approximations numériques près)
  - Implémenté efficacement dans de nombreux outils (PRISM, MRMC, GreatSPN)
  - Hypothèses probabilistes fortes
  - Taille de l'espace mémoire proportionnelle à la taille du processus stochastique
- Approche statistique
  - Logiques linéaires (basées sur LTL)
  - Intervalle de confiance : encadrement probabiliste
  - Taille de l'espace mémoire très faible
  - Facilement parallèlisable
  - Hypothèses probabilistes faibles
  - Inutilisable pour de très faibles probabilités (évènements rares)

Objectif

Développement d'une méthode pour traiter les évènements rares

Formalisation du problème



Une chaîne de Markov $\mathcal{C}$
Deux états absorbants s_-, s_+
atteints avec probabilité 1

Soit $\sigma = s \to s_1 \to s_2 \to \cdots \to s_\pm$
une trajectoire aléatoire dans $\mathcal{C}$

$$V_s = \begin{cases} 1 \text{ si } \sigma \text{ finit dans l'état } s_+ \\ 0 \text{ si } \sigma \text{ finit dans l'état } s_- \end{cases}$$

Objectif : Calculer $\mathbf{E}(V_{s_0})$ lorsque $\mathbf{E}(V_{s_0}) \ll 1$

Difficulté : $\mathbf{V}(V_{s_0})$ grand comparé à $\mathbf{E}^2(V_{s_0})$

Échantillonnage préférentiel

Principe : Substituer W_s à V_s de même espérance mais de variance réduite

1. Substituer $\mathbf{P}'$ à $\mathbf{P}$ telle que $\mathbf{P}(s,s') > 0 \Rightarrow \mathbf{P}'(s,s') > 0 \vee s' = s_-$
2. Pour chaque trajectoire $\sigma = s \rightarrow s_1 \rightarrow s_2 \cdots s_k \rightarrow s_\pm$
   On définit

$$W_s = \begin{cases} \frac{\mathbf{P}(s,s_1)}{\mathbf{P}'(s,s_1)} \cdot \frac{\mathbf{P}(s_1,s_2)}{\mathbf{P}'(s_1,s_2)} \cdot \ldots \cdot \frac{\mathbf{P}(s_k,s_+)}{\mathbf{P}'(s_k,s_+)} & \text{si } \sigma \text{ finit dans l'état } s_+ \\ 0 & \text{si } \sigma \text{ finit dans l'état } s_- \end{cases}$$

3. Estimer statistiquement $\mathbf{E}(W_{s_0})$

Échantillonnage préférentiel

Principe : Substituer W_s à V_s de même espérance mais de variance réduite

1. Substituer $\mathbf{P}'$ à $\mathbf{P}$ telle que $\mathbf{P}(s,s') > 0 \Rightarrow \mathbf{P}'(s,s') > 0 \vee s' = s_-$
2. Pour chaque trajectoire $\sigma = s \to s_1 \to s_2 \cdots s_k \to s_\pm$
   On définit

$$W_s = \begin{cases} \frac{\mathbf{P}(s,s_1)}{\mathbf{P}'(s,s_1)} \cdot \frac{\mathbf{P}(s_1,s_2)}{\mathbf{P}'(s_1,s_2)} \cdot \ldots \cdot \frac{\mathbf{P}(s_k,s_+)}{\mathbf{P}'(s_k,s_+)} & \text{si } \sigma \text{ finit dans l'état } s_+ \\ 0 & \text{si } \sigma \text{ finit dans l'état } s_- \end{cases}$$

3. Estimer statistiquement $\mathbf{E}(W_{s_0})$

La méthode est non biaisée

$$\forall s \in S,\ \mathbf{E}(W_s) = \mathbf{E}(V_s)$$

Objectif

$$\mathbf{V}(W_{s_0}) \ll \mathbf{V}(V_{s_0})$$

Échantillonnage préférentiel optimal

Un résultat « non opérationnel »

Il existe un échantillonnage préférentiel de variance nulle

Soit $\mu(s) = \mathbf{E}(V_s)$
Soit $\mathbf{P}'(s,t) = \frac{\mu(t)}{\mu(s)} \cdot \mathbf{P}(s,t)$

$$W_s = \frac{\mathbf{P}(s,s_1)}{\mathbf{P}'(s,s_1)} \cdot \frac{\mathbf{P}(s_1,s_2)}{\mathbf{P}'(s_1,s_2)} \cdots \frac{\mathbf{P}(s_k,s_+)}{\mathbf{P}'(s_k,s_+)} = \frac{\mu(s)}{\mu(s_1)} \cdot \frac{\mu(s_1)}{\mu(s_2)} \cdots \frac{\mu(s_k)}{1} = \mu(s)$$

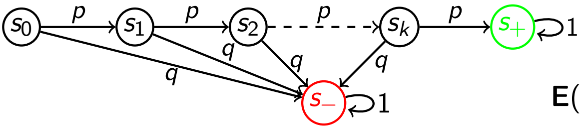
Donc $\mathbf{V}(W_s) = 0$

Problème

Nécessite de connaître μ ce que l'on cherche à estimer

Difficulté de mise en oeuvre

Cas simple (p' substitué à p)

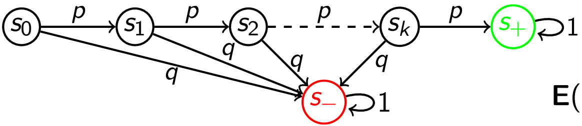


$$\mu(s_0) = p^{k+1}$$

$$\mathbf{E}(W_{s_0}) = \left(\frac{p}{p'}\right)^{k+1} p'^{k+1}$$

Difficulté de mise en oeuvre

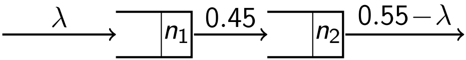
Cas simple (p' substitué à p)



$$\mu(s_0) = p^{k+1}$$

$$\mathbf{E}(W_{s_0}) = \left(\frac{p}{p'}\right)^{k+1} p'^{k+1}$$

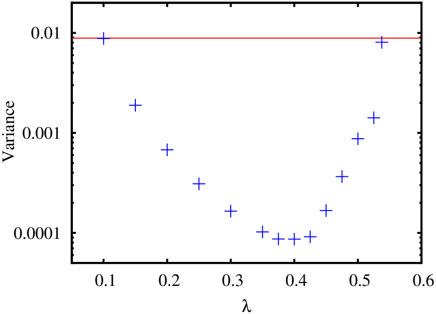
Cas plus complexe (λ' substitué à λ)



Atteindre 5 clients
versus
Revenir au repos

Pour le système originel ($\lambda = 0.1$)

$$\mathbf{E}(V_{s_0}) = 0.0089$$



État de l'art

Échantillonnage préférentiel asymptotiquement optimal
dans une classe d'échantillonnages
(P.Dupuis, A.D. Sezer, H. Wang 2007)

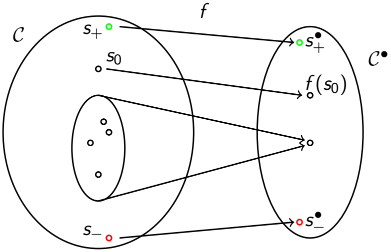
Utilisation d'heuristiques
(P.E Heegaard, W. Sandmann 2007)

Analyse au cas par cas
(Rubino, Tuffin 2009)

Aucune de ces méthodes ne fournit
un intervalle de confiance du résultat

Garantie de réduction de variance

Idée : substituer $\mu^\bullet(f(s))$ à $\mu(s)$ dans l'échantillonnage optimal



Quelle condition pour que W_{s_0} soit à valeurs dans $\{0; \mu^\bullet(f(s_0))\}$?

Une condition nécessaire

$\mathcal{C}^\bullet$ *est bornant :* $\mu(s_0) \leq \mu^\bullet(f(s_0))$

Le principe d'homogénéité conduit au renforcement suivant

$\mathcal{C}^\bullet$ *est uniformément bornant :* $\forall s \in S, \mu(s) \leq \mu^\bullet(f(s))$

Le principe de localité conduit à la condition suffisante suivante

$$\forall s \in S, \ \sum_{s' \in S} \mu^\bullet(f(s')) \cdot \mathbf{P}(s,s') \leq \mu^\bullet(f(s))$$

Dans ce cas, $(\mathcal{C}^\bullet, f)$ est une *réduction de $\mathcal{C}$ à variance garantie*

Échantillonnage préférentiel à variance garantie

Spécification

Pour tout $s' \neq s_-$, $\mathbf{P}'(s,s') = \frac{\mu^\bullet(f(s'))}{\mu^\bullet(f(s))}\mathbf{P}(s,s')$

$\mathbf{P}'(s,s_-) = 1 - \sum_{s' \neq s_-} \frac{\mu^\bullet(f(s'))}{\mu^\bullet(f(s))}\mathbf{P}(s,s')$ *(possible en vertu de la condition)*

Garantie de la réduction de la variance

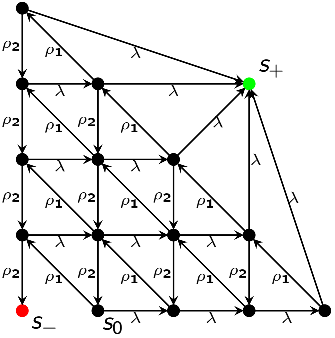
Soient $(\mathcal{C}^\bullet, f)$ une réduction à variance garantie de $\mathcal{C}$.
L'échantillonnage préférentiel précédent vérifie :

- Pour tout s, W_s est une variable aléatoire à valeurs dans $\{0, \mu^\bullet(f(s))\}$
- Par conséquent, $\mu(s) \leq \mu^\bullet(f(s))$
- Comme la distribution (itérée) suit une loi « binomiale », il est possible de calculer un intervalle de confiance.

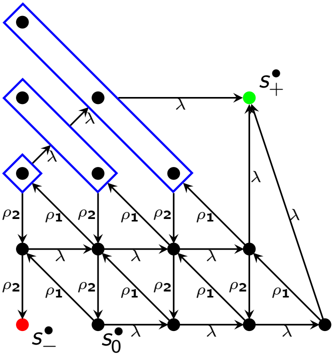
Exemple des files d'attente en tandem

On réduit le modèle en bornant la seconde file.

$$f(n_1, n_2) = \begin{cases} (n_1, n_2) & \text{si } n_2 \leq R \\ (n_1 + n_2 - R, R) & \text{sinon} \end{cases}$$



DTMC pour les files en tandem



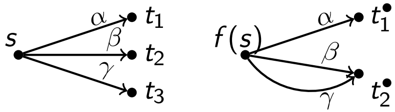
DTMC de la chaîne réduite par f

Comment vérifier la condition structurellement ?

On veut démontrer la condition de réduction sans analyse numérique.

$$\forall s \in S, \sum_{t \in S} \mu^\bullet(f(t)) \cdot \mathbf{P}(s,t) \leq \mu^\bullet(f(s))$$

Impliquée par deux conditions locales



et

$$\mu^\bullet(t_1^\bullet) \leq \mu^\bullet(f(t_1))$$
$$\mu^\bullet(t_2^\bullet) \leq \mu^\bullet(f(t_2))$$
$$\mu^\bullet(t_2^\bullet) \leq \mu^\bullet(f(t_3))$$

1. Mise en relation des transitions des deux chaines de mêmes probabilités *(examen des chaînes)*
2. Vérification de contraintes de la forme $\mu^\bullet(s) \leq \mu^\bullet(t)$ *(établissement d'un couplage approprié)*

Couplage

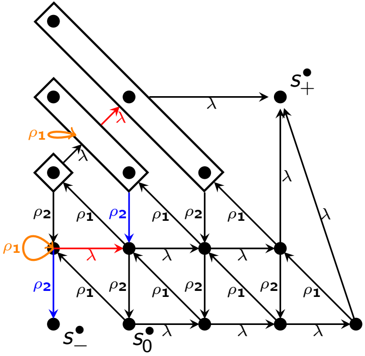
Un couplage de deux chaînes *(ici la même chaine)* est une chaîne sur un sous-ensemble de l'espace produit *(défini par R)* qui vérifie :

- que la chaîne produit projetée sur chaque composante est la chaîne correspondante.
- une contrainte additionnelle adaptée à la propriété à établir

1. $(n_1, n_2)\ R\ (n'_1, n'_2)$ ssi

$$\begin{cases} n_1 + n_2 \geq n'_1 + n'_2 \\ n_1 \geq n'_1 \end{cases}$$

2. $\forall s\ R\ s'\ s' = s_+ \Rightarrow s = s_+$.



Couplage

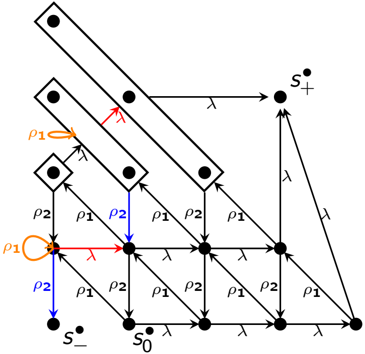
Un couplage de deux chaînes *(ici la même chaine)* est une chaîne sur un sous-ensemble de l'espace produit *(défini par R)* qui vérifie :

- que la chaîne produit projetée sur chaque composante est la chaîne correspondante.
- une contrainte additionnelle adaptée à la propriété à établir

1. $(n_1, n_2)\ R\ (n'_1, n'_2)$ ssi

$$\begin{cases} n_1 + n_2 \geq n'_1 + n'_2 \\ n_1 \geq n'_1 \end{cases}$$

2. $\forall s\ R\ s'\ s' = s_+ \Rightarrow s = s_+$.



$\boxed{\text{Propriété du couplage : } \forall s\ R\ s'\ \mu(s) \geq \mu(s')}$

Méthodologie

1. Spécifier un modèle réduit $\mathcal{M}^\bullet$ de chaîne de Markov associée $\mathcal{C}^\bullet$ et une fonction f.

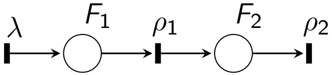
2. Etablir par examen de $\mathcal{C}$ et $\mathcal{C}^\bullet$ et par un couplage sur $\mathcal{C}^\bullet$ que la réduction est à variance garantie.

3. Calculer numériquement la fonction $\mu^\bullet$.

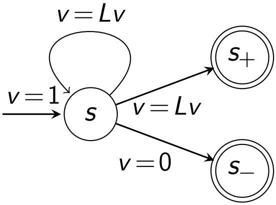
4. Calculer statistiquement $\mu(s_0)$ en utilisant l'échantillonnage préférentiel induit par $\mu^\bullet$.

COSMOS un outil de model checking statistique

- Modèle sous forme de réseaux de Petri à distributions générales



- Spécification de formules par un automate hybride

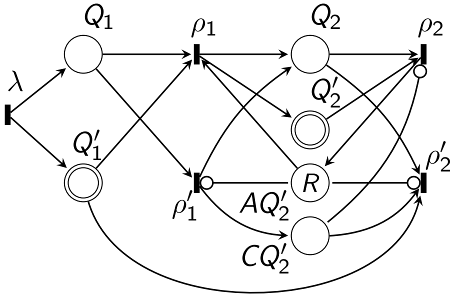


- Model checking et évaluation de performance

- Vitesse de simulation supérieure à PRISM

Adaptation de Cosmos

- Adaptation de la méthode aux modèles à temps continu
- Calcul de la valeur des trajectoires par l'automate hybride
- Gestion de l'image de f par le réseau de Petri
  à l'aide de places supplémentaires
- Modification des tirages probabilistes pour calculer $\mathbf{P}'$



Exemple du tandem

- $\lambda = 0.1$, $\rho_1 = \rho_2 = 0.45$,
- $N = 50$
- 20000 trajectoires
- $\mu(s_0) = 3.80122 \cdot 10^{-31}$

R	taille de $\mathcal{C}$	taille de $\mathcal{C}^\bullet$	$\mu^\bullet(s_0)$	$\mu(s_0)$ estimé	Intervalle de confiance	T (s) simulation
2	2500	100	1.24904E-28	3.96541E-31	2.25E-31	21.47
3	2500	150	2.28771E-30	3.78565E-31	2.76E-32	39.48
4	2500	200	6.55440E-31	3.80168E-31	9.63E-33	57.32
5	2500	250	5.10457E-31	3.79642E-31	4.18E-33	64.81
6	2500	300	3.97544E-31	3.80229E-31	1.86E-33	67.18
7	2500	350	3.97544E-31	3.79973E-31	8.90E-34	68.56

La taille de l'intervalle de confiance est satisfaisante même pour R petit.

N	R	taille de $\mathcal{C}$	$\mu(s_0)$ numérique	T (s) $\mathcal{C}$	T (s) $\mathcal{C}^\bullet$	taille de $\mathcal{C}^\bullet$	$\mu^\bullet(f(s_0))$	$\mu(s_0)$ estimé	Intervalle de confiance	T (s) simulation
100	30	10 201	0.01177	4	2	3131	0.017513	0.01156	8.6E-4	384
500	87	251 001	2.0619E-12	486	68	44088	3.139E-12	2.054e-12	2.13E-13	1031
1000	111	1E6	2.8694E-25	3880	389	112112	4.536E-25	2.824e-25	2.30E-26	1639
5000	150	25E6	#	#	13375	755151	2.265E-129	7.331e-130	1.10E-130	7920

On est capable d'estimer $\mu(s_0)$ avec un faible intervalle de confiance, pour de grandes valeurs de N.

Autres exemples

- Tandem (la seconde file est pleine, la première n'est pas bornée)
  - Système de taille infinie
  - avec un modèle réduit fini
- Tandem (la seconde file est pleine avant que la première soit pleine)
  - Garantie théorique
  - mais résultat expérimental peu satisfaisant
- Ruine parallèle
  - Système concurrent
  - avec construction du système réduit par désynchronisation
- Dîner des philosophes
  - Extension de la méthode sans garantie théorique
  - d'où un caractère « pathologique » de la distribution de W_{s_0}

Conclusion

- Contributions
  - Conception du premier échantillonnage préférentiel avec intervalle de confiance
  - Intégration dans un outil
  - Expérimentations variées
- Perspectives
  - Prise en charge plus générale des systèmes infinis
  - Extension aux distributions non markoviennes
  - (Semi-)Automatisation des preuves de couplages

Une partie de ce travail a donné lieu à une publication à MSR 2011