Terminaison en temps moyen fini de systèmes de règles probabilistes

Florent Garnier

DCS meeting
Col de porte, France

Mardi 10 juin 2008

france telecom

Loria

Objectifs

On veut définir un formalisme pour :

- Décrire des systèmes de transition probabilistes sur des espaces d'états infinis mais dénombrables.
- Exprimer des choix probabilistes et non-déterministes.
- Décrire des processus de décision Markoviens, chaînes de Markov etc ...

Objectifs

On désire avoir des méthodes de preuve de terminaison valables :

- Sur des systèmes de transitions probabilistes à espaces d'états infinis.
- Pour prouver la terminaison de protocoles probabilistes, tel que le CSMA/CA 802.11b.

Formalismes déjà existants

Formalismes de haut niveau modélisant Non Déterminisme et Probabilités :

- Automates Probabiliste [Segala,Lynch].
- Algèbres de processus probabilistes [Hannson,Johnson].
- Réseaux de Pétri probabilistes[Florin,Natkin].
- Réécriture et probabilités [Kumar,Sen,Meseguer,Agha],[Bournez,Kirchner].

Mais pas de méthodes de preuve de terminaison en temps moyen fini valable sur des systèmes à espace d'états infinis mais dénombrables modélisés par ces formalismes de haut niveau.

Plan de l'exposé

Nous allons présenter dans cet exposé :

1. Modèles :
   - Extension de la réécriture à la réécriture probabiliste.
2. Terminaison :
   - De critères impliquant la terminaison en temps moyen fini de ces systèmes.
3. La terminaison en temps moyen fini sous stratégies :
   - Critères généraux.
   - Conditions suffisantes.
4. Application : Le cas du CSMA/CA 802.11b WIFI.
5. Relations d'équivalences comportementales entre PRS.
6. Terminaison et équité.

Pourquoi choisir d'étendre la réécriture ?

La réécriture est un formalisme concis et puissant :

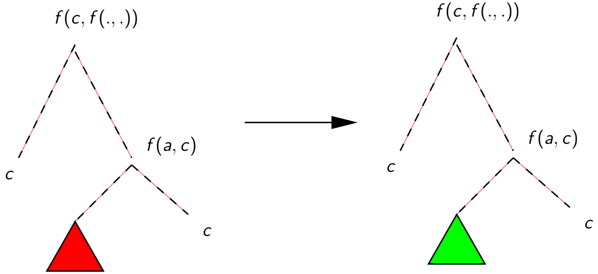
- $T(\Sigma, X)$ algèbre de termes, infinité d'états (les termes).
- Description concise d'une relation binaire sur $T(\Sigma, X) \times T(\Sigma, X)$ avec un ensemble fini de règles.
- Description de choix non déterministes.

La relation binaire sur les termes est appelée la relation de réduction.

La relation de réduction

Une règle

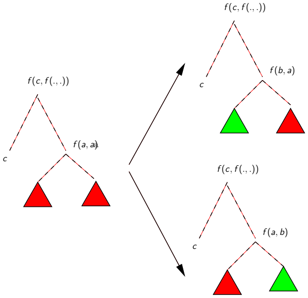
Relation induite



Réécriture et non-déterministe

Le non déterminisme a plusieurs sources :

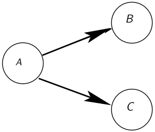
- Choix de la règle.
- Choix de la position.
- Choix de la substitution.



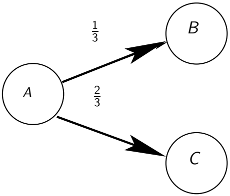
Problématique

La réécriture permet d'exprimer des transitions non déterministe.

$$\begin{array}{ccc} A & \longrightarrow & B \\ A & \longrightarrow & C \end{array}$$



Cependant, elle ne permet pas d'exprimer des transitions suivant une loi de probabilité.



Approches précédentes

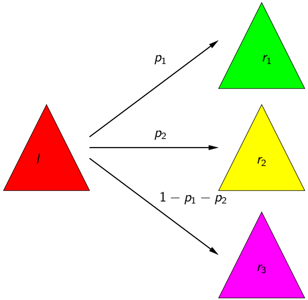
- Extension de la réécriture conditionnelle avec choix de successeurs probabiliste [Kumar,Sen,Meseger,Agha].
- Probabilistic Constraint Handling Rules [Früwirth,Di Pierro, Wiklicky].
- Appliquer des règles de réécriture suivant une loi de probabilité déterminée par une stratégie [Bournez,Kirchner,Hoyrupt][Bournez,Kirchner].

Nous décrivons les transitions probabilistes au niveau des règles.

Systèmes de réécriture probabilistes [Bournez,Garnier 05]

Définition (Règle de réécriture probabiliste)

Une règle de réécriture probabiliste est élément de $T(\Sigma, X) \times Dist(T(\Sigma, X))$.



Système de réécriture probabiliste

Définition (Système de réécriture probabiliste)

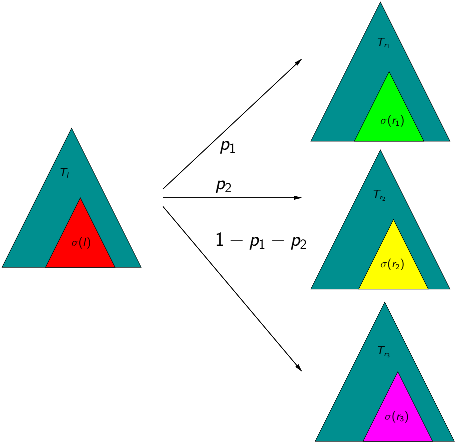
Un système de réécriture probabiliste est un ensemble fini de règles de réécriture probabilistes.

Remarque

Il faut redéfinir la nature de la relation de réécriture pour qu'elle mette en relation des termes avec des distributions de probabilité.

La relation de réduction probabiliste

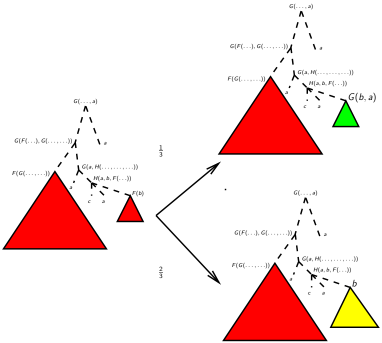
On construit la relation de réécriture $\rightarrow_{\mathcal{R}}$ d'un système de réécriture $\mathcal{R}$ de la façon suivante :



Non déterminisme et probabilités

$$r_1 : F(X) \rightarrow \begin{cases} G(X, a) & : \frac{1}{3} \\ b & : \frac{2}{3} \end{cases}$$

Choix Non déterministe puis choix probabiliste :



La question de la terminaison des systèmes de réécriture probabilistes

Que devient la notion de terminaison quand les transitions sont d'abord :

- Non déterministes (Choix de la règle, position de l'application, substitution).
- Puis probabilistes (Choix probabiliste d'un successeur).

Rappelons les différentes notions de terminaison dans le cas purement probabiliste.

Terminaison presque sûre des PRS

Les différentes terminaisons

$$r_1 \quad := \quad n+1 \rightarrow \begin{cases} n & : p_1 \\ n+2 & : 1-p_1 \end{cases}$$

$$\begin{array}{ll} p_1 < \frac{1}{2} & P(\neg Termine) > 0 \\ p_1 \geq \frac{1}{2} & P(Termine) = 1 \end{array}$$

Les différentes terminaisons

$$r_1 \quad := \quad n+1 \rightarrow \begin{cases} n & : p_1 \\ n+2 & : 1-p_1 \end{cases}$$

$$\begin{array}{lll} p_1 < \frac{1}{2} & P(\neg Termine) > 0 & \\ p_1 = \frac{1}{2} & P(Termine) = 1 & E[nbtransitions] = +\infty \\ p_1 > \frac{1}{2} & P(Termine) = 1 & E[nbtransitions] < +\infty \end{array}$$

Nous avons focalisé nos travaux pour l'étude de la terminaison en temps moyen fini.

La terminaison en temps moyen fini des PRS

Pour étudier la terminaison, on définit la notion de stratégie adaptée des :

- Politiques de PDM [Derman].
- Adversaires [Ségala, Lynch].
- Ordonnanceurs [Lehman Rabin, Vardi,Pnueli Zuck].

Remarque

Les stratégies ont pour but de résoudre les choix non déterministes.

Étude de la terminaison, définitions de base

Définition (Terme terminal)

Un terme t est terminal pour un système de réécriture probabiliste $\mathcal{R}$ s'il n'existe pas de $\mu \in Dist(T(\Sigma, X))$ tel que $t \to_{\mathcal{R}} \mu$.

Définition (histoire)

Une histoire de longueur $n+1$ est un séquence $a_0 \ldots a_n$ d'éléments de $T(\Sigma, X)$. Elle est non terminale si a_n ne l'est pas.

Définition (stratégie)

Une stratégie ϕ est une fonction des histoires non terminales vers les termes telle que $\phi(a_1 \ldots a_n) = \mu$ vérifie toujours $a_n \to_{\mathcal{R}} \mu$.

Quelques exemples de stratégies

Exemple

$$
\begin{aligned}
r_1 &:= & n+1 \rightarrow \begin{cases} n & : \frac{1}{2} \\ n+2 & : \frac{1}{2} \end{cases} \\
r_2 &:= & n+1 \rightarrow \begin{cases} n & : \frac{1}{3} \\ n+2 & : \frac{2}{3} \end{cases}
\end{aligned}
$$

On peut considérer les stratégies suivantes :

- ϕ_1 *qui applique toujours la règle* r_1*,*
- ϕ_2 *qui applique* r_1 *quand la longueur de l'histoire est paire et* r_2 *sinon.*
- ϕ_3 *qui applique* r_2 *si le dernier terme est paire.*

Terminaison presque sûre et presque sûre positive

Remarque

Les stratégies ont pour but de :

- *De rendre l'application des règles de réécriture déterministe.*
- *D'induire une mesure de probabilités sur les dérivation.*

Définition (Terminaison presque sûre positive (+ a.s))

Un système de réécriture termine positivement et presque sûrement (+a.s) si pour toute stratégie ϕ, tout terme a, la longueur moyenne d'une dérivation menant de a à un état terminal est fini.

Critères de terminaison en temps moyen fini

Définition (Variation d'une règle)

Soit $V : T(\Sigma, X) \to \mathbb{R}$, *et* $t \to_{\mathcal{R}} \mu$. *On note* $\Delta_\mu V(t)$ *le réel suivant :*

$$\Delta_\mu V(t) = \sum_{a \in T(\Sigma,X)} \mu(a)V(a) - V(t)$$

Exemple

$$g \to b : \tfrac{1}{3} | c : \tfrac{2}{3}$$
$$\Delta V(g) = \tfrac{1}{3} \times V(b) + \tfrac{2}{3} \times V(c) - V(g)$$

Critères de terminaison en temps moyen fini

On généralise les critères de terminaison au cas probabiliste:

Theorème (Soundness [*Bournez, Garnier*05])

Un système de réécriture probabiliste $\mathcal{R}$ termine en temps moyen fini s'il existe un $\epsilon > 0$, une fonction V inférieurement bornée telle que pour tout $t \in T(\Sigma, X)$ et $t \rightarrow_{\mathcal{R}} \mu$ la variation moyenne $\Delta_{\mu} V(t) < -\epsilon$.

Remarque

Ce résultat général n'est pas pratique à vérifier car il s'applique sur toute la relation de réécriture.

Conditions suffisantes pour la terminaison +a.s

Theorème ([*Bournez, Garnier*05])

Un système de réécriture $\mathcal{R}$ termine presque sûrement et positivement s'il existe une fonction $V : T(\Sigma, X) \rightarrow \mathbb{R}$ inférieurement bornée et un $\epsilon > 0$ tels que :

- *La variation moyenne de chaque règle est inférieur ou égal à $-\epsilon$.*
- *V a sa décroissance préservée par contexte et par substitution.*

Exemple d'application

Exemple

Si $p_{1_1} + p_{1_2} < p_{1_3}$ *et* $2 * p_{2_1} + p_{2_3} < p_{2_2}$
Alors le système de réécriture suivant

$$\begin{array}{lcl} f(x) & \rightarrow & f(f(x)) : p_{1_1} | g(f(x)) : p_{1_2} | x : p_{1_3} \\ f(h(f(x), x)) & \rightarrow & h(g(f(f(x))), f(x)) : p_{2_1} \\ & & g(f(x)) : p_{2_2} | f(g(f(f(x)))) : p_{2_3} \end{array}$$

termine $+$*a.s.* .

Preuve de la terminaison de l'exemple

Preuve

Soit V la fonction qui compte le nombre de symboles de fonction d'un terme. Calculons le drift de la première règle :

$$\begin{aligned}\Delta_1 &= V(f(f(x)) \times p_{1_1} + V(g(f(x))) \times p_{1_2} \\ &\quad + V(x) \times p_{1_3} - V(f(x)) \\ \Delta_1 &= p_{1_1} + p_{1_2} - p_{1_3}\end{aligned}$$

Ainsi que celui de la seconde :

$$\Delta_2 = 2 \times p_{2_1} + p_{2_3} - p_{2_2}$$

Si $\Delta_1 < 0$ et $\Delta_2 < 0$ alors le PRS est $+$a.s. terminant, car la décroissance de V est préservée par contexte et substitution.

Terminaison presque sûre sous stratégies

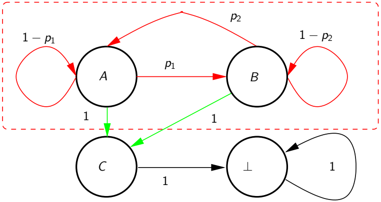
Terminaison sous stratégies

Motivation :

- Programme = Règles + Stratégies.
- Terminaison +a.s difficile à calculer.
- On s'intéresse souvent qu'à un sous ensemble des dérivations.
- Permet de ne considérer que certains évènements clef.

Exemple de terminaison sous stratégies

$$
\begin{array}{lcl}
A & \rightarrow & \begin{cases} B & : p_1 \\ A & : 1 - p_1 \end{cases} \\
B & \rightarrow & \begin{cases} A & : p_2 \\ B & : 1 - p_2 \end{cases} \\
A & \rightarrow & \{C : 1 \\
B & \rightarrow & \{C : 1
\end{array}
$$



Évaluation des dérivations sur le long cours

- L'évaluation d'une dérivation peut augmenter pourvu que sa moyenne décroisse au bout d'un certain temps.
- On peut envisager d'étudier la terminaison de systémes dont les règles ont une variation moyenne positive ou nulle.
- Identifier les stratégies qui sélectionnent les bonnes règles assez souvent.
- On étudie l'évaluation des dérivations par V sur le long cours pour prouver la terminaison sous stratégies.

Espérance de l'évaluation d'une dérivation

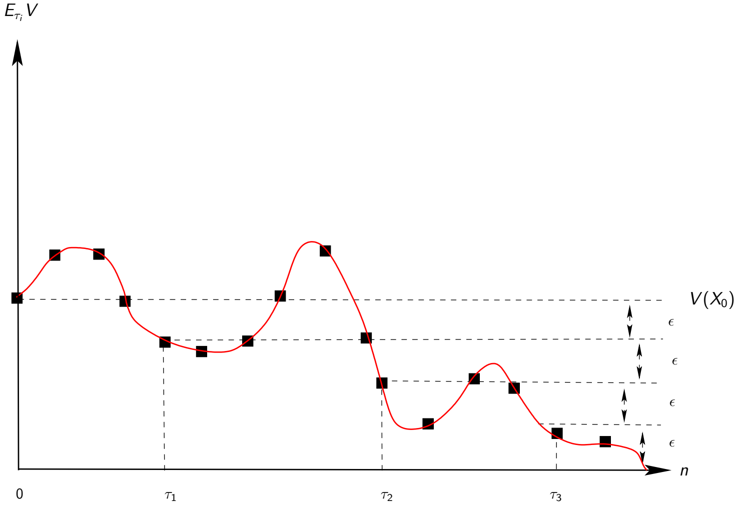
Définition (Espérance de V au temps τ)

Soit $(\pi_n)_{n\in\mathbb{N}}$ une dérivation sous la stratégie ϕ, $V : T(\Sigma, X) \to \mathbb{R}$ et $\tau \in \mathbb{N} \cup \{\infty\}$ un temps d'arrêt, presque sûrement fini.
On note $E_\tau V$ l'espérance V au temps τ :

$$E_\tau V = E[V(\pi_\tau)]$$

quand cette valeur existe.

Décroissance sur le long cours



Probabilité d'événements

Définition (Temps moyen de sélection borné)

Soit D un ensemble de règles. Une classe de stratégies Φ a un temps de sélection moyen borné $\alpha \in \mathbb{R}$ pour D, si $\forall \phi \in \Phi$, $\forall h$ histoire, le nombre moyen de règles déclenchées pour atteindre un état terminal ou une règle de D est majoré par α.

Terminaison positive et presque sûre sous stratégies

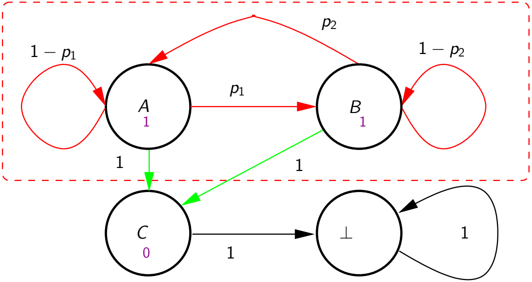
Theorème (Terminaison +a.s sous stratégies [*Bournez*, *Garnier*06])

Un système de réécriture $(T(\Sigma, X), \rightarrow_{\mathcal{R}})$ termine +a.s. sous Φ s'il existe :

- *Une fonction $V : T(\Sigma, X) \rightarrow \mathbb{R}$, avec $\inf_{t \in T(\Sigma,X)} V(t) > -\infty$*
- *un ensemble D de règles tel que D a un temps de sélection moyen borné α sous les stratégies ϕ de Φ pour un α positif fixé.*
- *un réel $\epsilon > 0$ tel que $\forall \phi \in \Phi$ $\forall h = a_0, \ldots, a_n$*

$$E_{\tau_{D,h}} V \leq V(a_n) - \epsilon$$

Exemple d'application



Soit ϕ la stratégie qui déclenche la règle provocant la transition représentée en vert quand les deux dernier termes atteints sont B. Le PRS termine sous +a.s. sous cette stratégie.

Le CSMA/CA en quelques mots

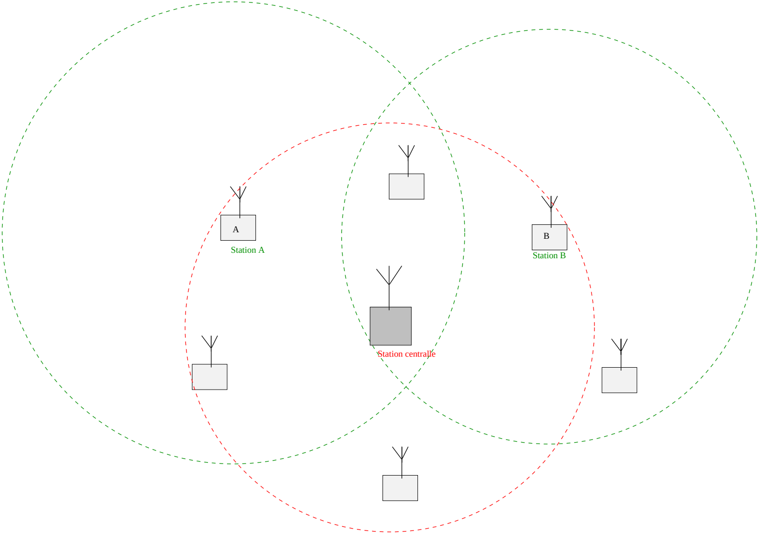
- Protocole utilise pour les réseaux WIFI.
- Protocole probabiliste distribué.
- Sert à réserver la bande passante et à transmettre des message.

Question : Est-ce que toutes les stations peuvent émettre un message au bout d'un temps moyen fini ? Existe-t-il des configurations de blocage mutuel ?
Réponse :

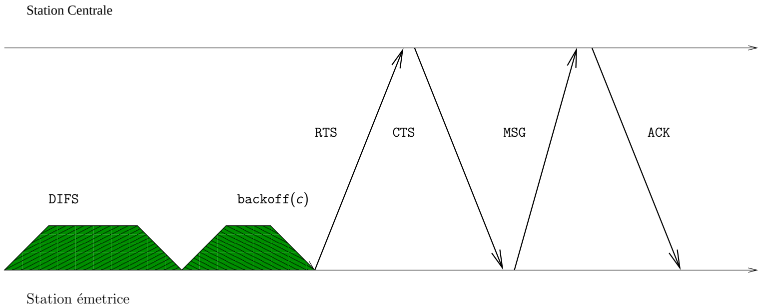
Theorème

Le protocole se comporte bien.

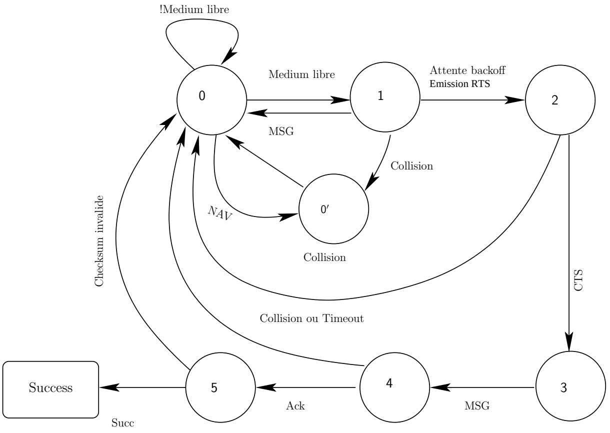
Un réseau WIFI centralisé



Envoi d'un message



Phases de l'algo du CSMA/CA



Représentation du système et du protocole

Les termes représentent l'état des différentes parties du système.

- Un terme pour représenter les stations clientes.
- Un terme pour représenter la station centrale.
- Un terme pour représenter l'ensemble du système.

Les règles de réécriture permettent de simuler l'évolution du système. Deux types de règles :

- Modification des automates.
- Règles de tri, mise à jour du terme de système.

Une stratégie applique les règles de réécriture dans l'ordre qui correspond à l'exécution de la simulation du CSMA.

Preuve de la terminaison sous stratégie

- On construit une fonction V dont le drift est négatif sur les règles modifiant les automates.
- Le drift de V est nul pour toutes les autres règles.
- Les premières règles ont un temps moyen de sélection borné.

On applique le théorème sur la terminaison en temps moyen fini sous stratégie.

Relation d'équivalence comportementale

La relation de bisimulation probabiliste sur les PRS :

- Adaptation de la Probabilistic Bisimulation [Segala, Lynch]
- Conserve les propriétés quantitatives et qualitatives sur les formules d'états invariantes sur les classes d'équvalence.
- Préserve la terminaison A.S sûre et +AS.

Relation d'équivalence comportementale

Utiliser de telles relation permet de :

- Montrer que des PRS sont bisimilaires à des systèmes terminants.
- Passer d'un espace d'états infinis, à un ensemble de classes d'équivalence finies.

Etude de la terminaison sous strategies équitables

- Sur le graph des paires de dépendances des TRS classiques.
- L'équité simple/forte n'es pas assez fine pour garantir une terminaison en temps moyen fini.
- Extension aux opérateurs de fréquence.

Conclusion

Le formalisme de réécriture probabiliste permet :

- De représenter des systèmes de transition probabiliste sur des espaces d'état infinis,
- D'étudier leur terminaison en temps moyen fini,
- D'étudier leur terminaison en fonction du contrôle de l'application des règles,
- De modéliser des systèmes probabilistes complexes et étudier leur terminaison.

La terminaison sous stratégie permet :

- D'étudier la terminaison en temps moyen fini protocoles probabilistes.
- De déterminer plus facilement des fonctions certifiant la terminaison presque sûre positive sous stratégies.

Perspectives

- Comparer l'expressivité des systèmes de réécritures probabilistes.
- Développer un formalisme d'opérateurs de fréquence d'évènements.
- Automatisation de la preuve de terminaison en utilisant RPO.
- Intégrer ces stratégies dans des langages déjà existant, tel TOM.