

# Processus généralisé de décision markovien partiellement observable utilisant les arbres de décisions

Rafic Faddoul, Wassim Raphael, Alaa Chateaneuf, Abdel-Hamid Soubra

► **To cite this version:**

Rafic Faddoul, Wassim Raphael, Alaa Chateaneuf, Abdel-Hamid Soubra. Processus généralisé de décision markovien partiellement observable utilisant les arbres de décisions. 41èmes Journées de Statistique, SFdS, Bordeaux, 2009, Bordeaux, France, France. 2009. <inria-00386770>

**HAL Id: inria-00386770**

**<https://hal.inria.fr/inria-00386770>**

Submitted on 22 May 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Processus généralisé de décision markovien partiellement observable utilisant les arbres de décisions

Rafic Faddoul<sup>a,b</sup> Wassim Raphael<sup>a,1</sup> Alaa Chateaneuf<sup>c</sup> Abdel-Hamid SOUBRA<sup>b</sup>

<sup>a</sup> ESIB, Université Saint-Joseph, Mar Roukos, BP 11-154, Riad El Solh, Beyrouth, Liban.

<sup>b</sup> Institut de Recherche en Génie Civil et Mécanique, Université de Nantes, BP 152, 44603 St-Nazaire Cedex, France.

<sup>c</sup> LGC – UBP, Polytech’Clermont-Ferrand, Campus des Cézeaux, 63174, Aubière, France

## Résumé

Cet article propose un processus généralisé de décision markovien partiellement observable (GPOMDP) combinant l’analyse de décisions à la programmation dynamique. Le modèle est appliqué à un problème d’optimisation de la maintenance. La méthodologie proposée permet une modélisation mathématique simple pour l’optimisation de séquences de décisions complexes destinées à être appliquées durant chaque étape.

## Abstract

This paper proposes a Generalized Partially Observable Markov Decision Process (GPOMDP) combining decision analysis and dynamic programming. The model is applied to a maintenance optimization problem. The proposed methodology allows an easy mathematical modeling for optimizing complex sequences of decisions that are to be undertaken during each stage.

*Mots clés : processus de décisions markoviens, programmation dynamique, analyse de décisions, optimisation, maintenance.*

## 1 Introduction

Depuis que la programmation dynamique fut proposée par Abraham Wald, au milieu des années quarante comme solution pour les problèmes d’optimisation des décisions séquentielles, elle fut largement utilisée dans les problèmes de planification de la maintenance. Cependant, cette méthodologie avait une limitation significative, à savoir, une inspection parfaite était implicitement supposée être effectuée au début de chaque étape. Evidemment, dans beaucoup de cas, une telle supposition s’avère être irréaliste. Une évolution dans ce sens fut réalisée par les processus de décisions markoviens partiellement observables (*i.e. Partially Observable Markov Decision Process POMDP*) (Drake 1962, Eckles 1968, Monahan 1982, Ben Akiva et al. 1991) dans lesquels il est présumé que l’état du système connu par le gestionnaire au début de chaque étape est incertain. Une extension importante des POMDP permettant d’optimiser au début de chaque étape une séquence d’actions composée d’une inspection choisie parmi plusieurs alternatives suivie d’une action d’entretien choisie parmi plusieurs alternatives fut proposée (Madanat 1991, Smilowitz et Madanat 2000, Corotis et al. 2005).

Dans le présent article, nous proposons une généralisation des POMDP, dans laquelle une séquence d’actions de n’importe quelle complexité durant chaque étape peut être optimisée dans le cadre de la programmation dynamique. Par exemple, une séquence de décisions composée de deux inspections séquentielles suivies d’une action d’entretien est présentée. Dans un problème

---

<sup>1</sup> Corresponding author: wassim.rafael@fi.usj.edu.lb  
Emails: faddoul\_rafic@yahoo.com

d'optimisation de la maintenance, ce type de modélisation est utile quand par exemple une inspection plus coûteuse et plus précise est exécutée en se basant sur les résultats d'une inspection relativement bon marché ou quand des inspections spécialisées pour détecter certains états du système sont exécutées en se basant sur les résultats d'une inspection qui a une précision plus ou moins uniforme sur tout l'univers des états possibles du système. Par souci de concision, on désignera le modèle proposé dans la suite de cet article par GPOMDP (i.e. Generalized Partially Observable Markov Decision Process), ce qui signifie qu'il s'agit d'un Processus Généralisé de Décisions Markovien Partiellement Observable. Cette méthodologie est décrite dans le contexte d'un problème d'optimisation de la maintenance, quoique, l'étendue de son utilité dépasse largement cette classe particulière de problèmes.

## 2 Processus Généralisé de Décisions Markovien Partiellement Observable

### 2.1 Hypothèses de base

1. La variable  $\theta^n \in \Theta = \{\theta_1, \theta_2, \dots, \theta_m\}$  qui décrit l'état de la structure au début de l'étape  $n$ , prend ses valeurs d'un ensemble fini et dénombrable.
2. La structure est soumise à un processus Markovien stochastique de dégradation.
3. Au début de chaque étape, le gestionnaire du système doit choisir une séquence d'actions à appliquer. Cette séquence sera composée généralement d'une ou de plusieurs inspections et/ou actions d'entretien séquentielles.
4. Les techniques d'inspection  $i \in I = \{i_0, i_1, \dots, i_p\}$  seront choisies parmi un ensemble fini d'alternatives. Les résultats de ces inspections sont incertains mais on peut les décrire par des distributions de probabilité.
5. Les actions d'entretien  $a^n \in A = \{a_0, a_1, \dots, a_a\}$  seront choisies parmi un ensemble fini d'alternatives.
6. Etant donné l'état  $\theta^n$  de la structure au début de l'étape  $n$ , l'effet d'une action d'entretien est incertain, on supposera que cette incertitude peut être décrite par une distribution de probabilité qui sera exprimée par une matrice de transition carrée  $A_{an}$ . En d'autres termes, l'état  ${}^a\theta^n$  de la structure durant l'étape  $n$ , après l'application de l'action  $a^n$ , est une variable aléatoire. Ainsi, la matrice  $M_i$  de transition de  $\theta^n$  à  $\theta^{n+1}$  sera le produit de la matrice de transition  $A_{an}$  de  $\theta^n$  à  ${}^a\theta^n$  par la matrice de transition Markovienne  $M$  de  ${}^a\theta^n$  à  $\theta^{n+1}$  i.e  $M_i = A_{an} \times M$ .
7. Les coûts considérés sont :
  - $ci(i)$  : coût de l'application de la technique d'inspection  $i$ .
  - $ca(a)$  : coût de l'action  $a$ .
  - $cs({}^a\theta^n)$  : coût engendré par le fait que la structure se trouve dans l'état  ${}^a\theta^n$  durant l'étape  $n$ , suite à l'application de l'action d'entretien.
8. Le critère d'optimalité adopté est la maximisation de l'utilité espérée  $u$  (Von Neumann & Morgenstern, 1944) qui est définie comme étant l'inverse du coût espéré total, i.e.  $u = -c$ .

### 2.2 Programmation dynamique

Dans un POMDP, l'« état de croyance » de la structure au début de l'étape  $n$  peut être défini par un vecteur  $v^n = [v_1^n, v_2^n, \dots, v_m^n]$  ou les  $v_i^n$  sont les probabilités associées avec les différents états de dégradation  $\theta_i$ , i.e.  $v^n = [\text{Pr}(\theta_1^n), \text{Pr}(\theta_2^n), \dots, \text{Pr}(\theta_m^n)]$ . L'état de croyance  $v^{n+1}$  de la structure au début de l'étape  $n+1$  sera égal au produit matriciel du vecteur  $v^n$  par la matrice de transition  $A_{an}$ , relative à l'action probabiliste  $a^n$  puis par la matrice de transition  $M$  relative au processus de

dégradation Markovien; i.e.  $v^{n+1} = v^n \times A_{an} \times M$ . Le vecteur  $v^{n+1}$  peut être éventuellement mis à jour sachant les résultats d'une inspection. En outre, sachant l'état de croyance de la structure au début de chaque étape, la solution d'un POMDP classique prescrira les actions d'entretien optimales qui devront être exécutées.

Cependant, dans un tel POMDP, une unique technique d'inspection imparfaite est supposée être appliquée au début de chaque étape durant toute la vie escomptée de l'ouvrage. Donc, une optimisation de la planification des inspections n'était pas possible i.e. un choix optimal, au début de chaque étape, de la technologie d'inspection(s) parmi un ensemble d'alternatives possibles n'était pas possible. Pour pallier à ce défaut, nous proposons un POMDP généralisé où le gestionnaire de la structure a la possibilité de choisir de manière optimale une séquence d'actions de n'importe quelle complexité destinée à être appliquée à la structure durant chaque étape. Une telle séquence d'actions est formée d'une ou de plusieurs inspections et/ou actions séquentielles. Au début de chaque étape, sachant l'état de croyance de la structure, le problème de détermination de la séquence optimale est très similaire, nonobstant quelques modifications nécessaires, à un problème de décision face à l'incertitude tel qu'il est appréhendé par l'analyse de décisions (arbre de décisions) (Simar 2002). Par conséquent, on propose d'utiliser l'analyse de décisions comme relation récursive de la programmation dynamique.

### 2.3 Analyse de décisions

On présente dans cette section, comme exemple, les calculs d'analyse de décisions destinés à être utilisés comme relation récursive du GPOMDP dans le cas particulier de l'optimisation d'une séquence d'actions composée de deux inspections séquentielles imparfaites suivies d'une action d'entretien imparfaite. Un exemple d'un tel sous-modèle est illustré dans la figure 1.

Ayant saisi la logique sous-jacente à un tel sous-modèle, le lecteur sera capable de concevoir d'autres sous-modèles adaptés à d'autres problèmes particuliers. A l'instar de n'importe quelle relation récursive de programmation

dynamique, ce sous-modèle utilisera comme données d'entrée les coûts

optimaux calculés pour les différents états possibles de la structure au début de l'étape  $n + 1$  pour calculer le coût optimal associé à un état de croyance donné de la structure au début de l'étape  $n$ .

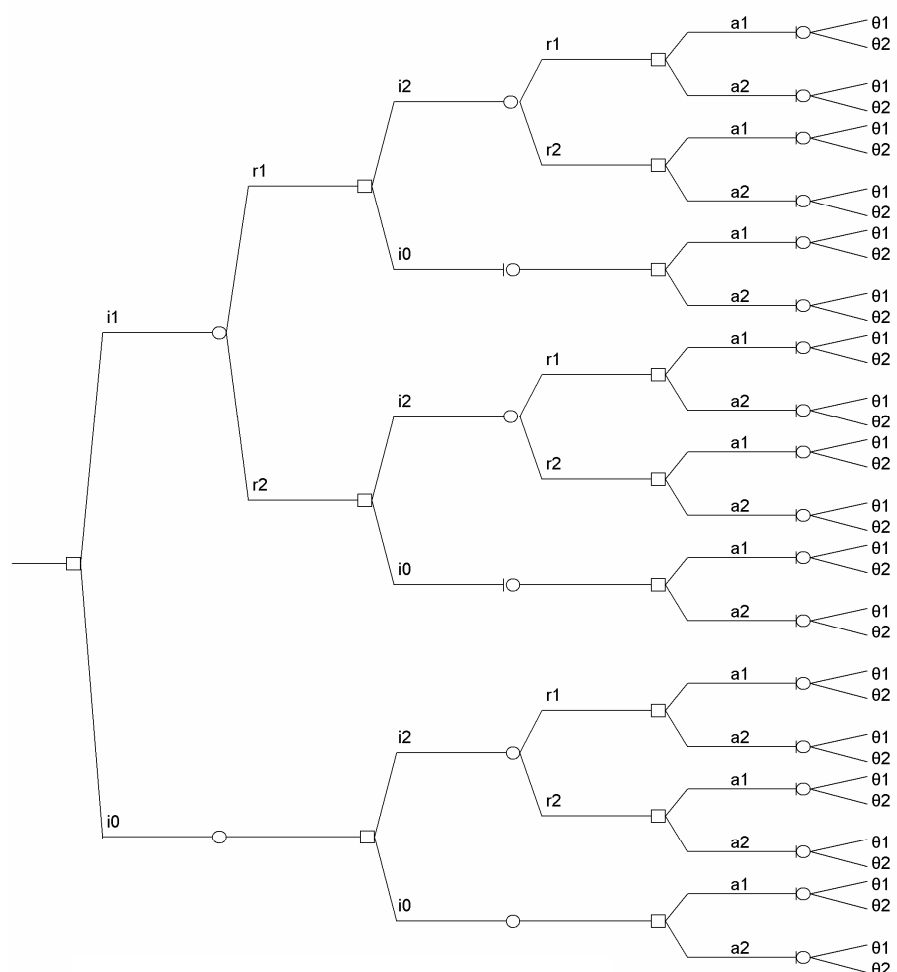


Figure 1 - Exemple d'arbre de décisions

Afin de pouvoir intégrer les sous-modèles d'analyse de décisions dans le cadre d'un POMDP, quelques modifications aux calculs usuels utilisés dans une analyse de décisions sont nécessaires afin de prendre en compte, d'une part, l'action simultanée de deux actions stochastiques, à savoir, l'action d'entretien et l'effet de la détérioration et d'autre part, les coûts optimaux calculés pour les différents états de croyance possibles de la structure au début de l'étape  $n + 1$ , pour déterminer les coûts optimaux des différents états de croyance de la structure au début de l'étape  $n$ .

Dans ce qui suit, on notera  $c(v^n)_{|x,y,\dots}$  les coûts espérés conditionnels. On supposera le cas général où les inspections sont imparfaites, i.e. le résultat d'une inspection, sachant le vrai état de la structure  $\theta^n$  et la technologie  $i$  de l'inspection, sera caractérisé par une distribution de probabilité conditionnelle ( $\Pr[r_1|\theta^n, i], \Pr[r_2|\theta^n, i], \dots, \Pr[r_m|\theta^n, i]$ ).

Sachant l'état de croyance  $v^n$  et la technique d'inspection  $i$ , la probabilité d'obtenir les différents résultats  $r$  peut être calculée par :

$$\Pr[r] = \Pr[r|\theta_1^n, i] \times \Pr[\theta_1^n] + \Pr[r|\theta_2^n, i] \times \Pr[\theta_2^n] + \dots + \Pr[r|\theta_m^n, i] \times \Pr[\theta_m^n] \quad (1)$$

D'autre part, sachant l'état de croyance  $v^n$ , la technique d'inspection  $i$  et le résultat  $r$  de l'inspection, l'état de croyance bayésien *a posteriori* ' $v^n$ ' peut être obtenu en calculant chacun de ses composants par :

$$\Pr[\theta_k^n | r, i] = \frac{\Pr[r | \theta_k^n, i] \times \Pr[\theta_k^n]}{\Pr[r | \theta_1^n, i] \times \Pr[\theta_1^n] + \Pr[r | \theta_2^n, i] \times \Pr[\theta_2^n] + \dots + \Pr[r | \theta_m^n, i] \times \Pr[\theta_m^n]} \quad (2)$$

Le coût total,  $c(v^n)_{|i^1, r^1, i^2, r^2, a^n, \theta^n}$ , dépendra de la technique  $i^1$  de la première inspection et de son résultat  $r^1$ , de la technique  $i^2$  de la seconde inspection et de son résultat  $r^2$ , ainsi que de l'action  $a^n$  et de l'état de la structure  $\theta^n$  :

$$\begin{aligned} c(v^n)_{|i^1, r^1, i^2, r^2, a^n, \theta^n} &= ci^1(i) + ci^2(i) + ca(a^n) + \frac{1}{1+\alpha} \times^* c((v^{n+1}|i^1, r^1, i^2, r^2, a^n, v^n)) + \mathbf{E}_{a^n|\theta^n, a^n} [cs(a^n)] \\ &= ci^1(i) + ci^2(i) + ca(a^n) + \frac{1}{1+\alpha} \times^* c(v^n \times A_{an} \times M) + \sum_k cs(a^n \theta_k^n) \times a_{jk}^{an} \end{aligned} \quad (3)$$

où :

- $a_{jk}^{an}$  est l'élément de la ligne  $j$  et de la colonne  $k$  de la matrice de transition  $A_{an}$  relative à l'action  $a^n$
- $\alpha$  est le taux d'escompte

L'état de croyance bayésien *a posteriori* ' $v^n$ ' sera calculé en appliquant la formule de Bayes (2) à deux reprises sachant  $i^1, r^1, i^2$  et  $r^2$ .

Le coût espéré optimal  $c(v^n)$  sera calculé en déterminant tour à tour les décisions  $i^1, i^2$  et  $a^n$  qui minimisent  $c(v^n)_{|i^1, r^1, i^2, r^2, a^n}$  et les espérances par rapport aux variables aléatoires  $r^1, r^2$  et  $\theta^n$ .

On pourra spécifier un type d'inspection type  $i_0$  et un type d'action  $a_0$  ayant des coûts nuls et qui consisteront à ne rien faire.

Les décisions optimales déterminées ne sont pas nécessairement uniques. Le cas échéant, le choix entre deux décisions équivalentes serait arbitraire, s'il n'existe pas d'autres considérations préférentielles.

### 3 Application numérique

La planification de l'inspection, de la maintenance et de la réhabilitation d'un pont en béton armé a été réalisée à l'aide d'un logiciel d'optimisation utilisant le processus généralisé de décisions markovien GPOMDP que nous avons développé. Trois séquences de décisions ont été considérées : (i) choix optimal d'une action de maintenance, (ii) choix optimal d'une inspection suivie d'une action et enfin, (iii) choix optimal de deux inspections séquentielles suivies d'une action. Les coûts espérés optimaux pour des horizons de planification allant de deux à vingt années ont été calculés (figure 2). On remarque que les coûts espérés à long terme résultants du plan inspection-inspection-action ont été les plus faibles sur toute l'étendue des horizons considérés.

On constate par exemple que le gain d'optimalité créé par l'utilisation d'un plan inspection-inspection-action s'élève à 5% par rapport à un plan inspection-action.

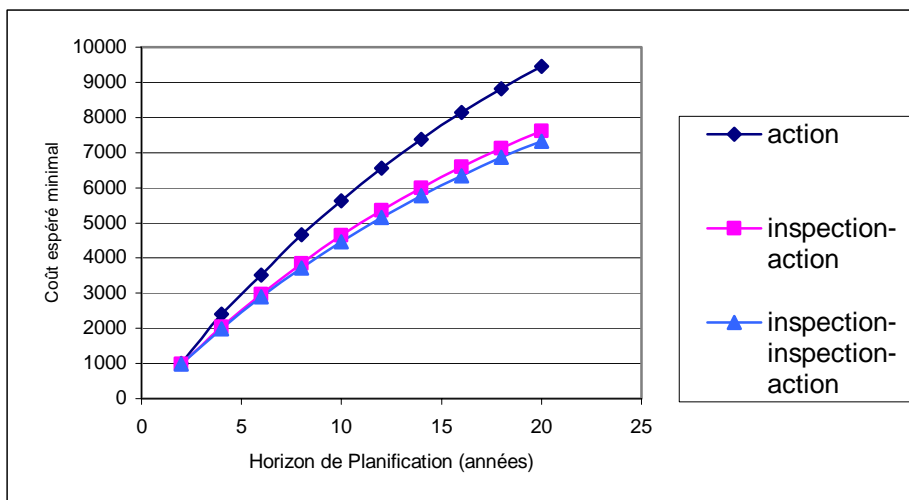


Figure 2 - Coût espéré minimal pour trois plans de maintenance différents

Les stratégies optimales d'inspection, de maintenance et de réhabilitation pour chaque état de croyance durant chaque étape seront prescrites par la solution du GPOMDP. Si pour une raison quelconque (d'ordre politique, difficultés logistiques, budget limité durant une certaine étape, etc.), le gestionnaire du pont était contraint d'adopter des décisions non optimales à un temps donné, la solution fournie par notre modèle indiquera dynamiquement les décisions optimales à prendre dorénavant en se basant sur l'état de croyance survenu sans avoir besoin de recalculer l'algorithme.

Comme exemple, le coût espéré minimal ainsi que les coûts directs espérés qui doivent être payés par le gestionnaire durant chaque étape pour la maintenance de la structure, sont présentés dans le tableau 1 sous la forme

Tableau 1. Coûts espérés pour des choix différents de techniques d'inspection à être appliquées au début de la 1<sup>ère</sup> année.

	$i_0$	$i_1$	$i_2$	$i_3$
coût espéré total	7811.50	7747.11	7607.80	7630.54
coût direct (année 1)	0	456	412	652
coût direct (année 2)	707.68	573.47	607.22	477.42
coût direct (année 4)	574.83	550.74	520.73	523.67
coût direct (année 6)	538.72	531.84	529.38	523.38
coût direct (année 8)	530.51	528.18	526.27	525.73
coût direct (année 10)	506.89	506.24	506.01	505.66
coût direct (année 12)	511.27	511.17	511.06	511.01
coût direct (année 14)	467.19	467.15	467.13	467.12
coût direct (année 16)	468.46	468.44	468.44	468.43
coût direct (année 18)	242.52	242.52	242.52	242.52

d'une analyse de simulation de scénario (what if) qui examine la sensibilité des coûts espérés dans l'éventualité que le gestionnaire choisit une technique d'inspection non optimale au début de la première étape.

## 4 Conclusion

Dans cet article, un processus généralisé de décision markovien partiellement observable (GPOMDP) est proposé en combinant l'analyse de décisions à la programmation dynamique. Dans le modèle proposé, l'état de la structure au début de chaque étape est défini en tant qu'état de croyance. Par conséquent, des sous-modèles d'analyse de décisions, pouvant être adaptés pour répondre aux exigences spécifiques de problèmes particuliers, servent d'outils convenables pour l'optimisation de séquences de décisions pour chaque étape. Nous avons illustré notre méthodologie en établissant les équations mathématiques dans le cas particulier d'une séquence de décisions composée de deux inspections séquentielles suivies d'une action d'entretien. Une application numérique a été présentée comprenant une analyse de sensibilité pour des horizons de planification s'étendant de deux à vingt ans. Une analyse de scénario (what if) a été menée pour déterminer l'influence du choix d'une décision non optimale sur les coûts espérés. Il a été montré qu'en permettant au gestionnaire d'un processus de décision markovien partiellement observable de planifier la mise en œuvre d'inspections imparfaites, cela permet de réaliser une augmentation significative de l'utilité espérée.

## 5 Références

- [1] Ben Akiva, M., Humplick, F., Madanat, S. and Ramaswamy, R., 1991. Latent performance approach to infrastructure management. *Transport. Res. Record*, **1311**, 188-195.
- [2] Corotis, R.B., Ellis, J.H. and Jiang, M., 2005. Modeling of risk-based inspection, maintenance and life-cycle cost with partially observable Markov decision process. *Structure and Infrastructure Engineering*, **1**, 75-84.
- [3] Drake, A., 1962. Observation of a Markov process through a noisy channel. Thesis (Sc.D.), Massachusetts Institute of Technology, Cambridge.
- [4] Eckles, J.E., 1968. Optimum maintenance with incomplete information. *Operations Research*, **16**, 1058-1067.
- [5] Faddoul, R., Raphael, W., Chateauneuf, A., 2009. A Generalized Partially Observable Markov Decision Process updated by decision trees for maintenance optimization, *Journal of Structure and Infrastructure Engineering*, Volume XX.
- [6] Madanat, S.M., 1991. Optimization sequential decisions under measurement and forecasting uncertainty: application to infrastructure inspection, maintenance and rehabilitation. Thesis (PhD). Massachusetts Institute of Technology.
- [7] Monahan, G.E., 1982. A Survey of partially observable Markov decision processes; theory, models, and algorithms. *Management Science*, **28**, 1-16.
- [8] Simar, L., 2002. Le Paradigme Bayésien. In Dreesbeke, J. J., Fine, J., Saporta, G., eds. *Méthodes Bayésiennes en statistique*. Paris: Société française de statistique, 19-60.
- [9] Smilowitz, K. and Madanat, S., 2000. Optimal Inspection and Maintenance Policies for Infrastructure Structures Under Measurement and Prediction Uncertainty. University of California, Berkeley, in *TRB (Transportation Research Circular)*, **498**.
- [10] von Neumann, J., and Morgenstern, O., 1944. Theory of games and economic behavior. Princeton University press.