Cours 3 et 4

Processus décisionnels Markoviens

Patrice PERNY

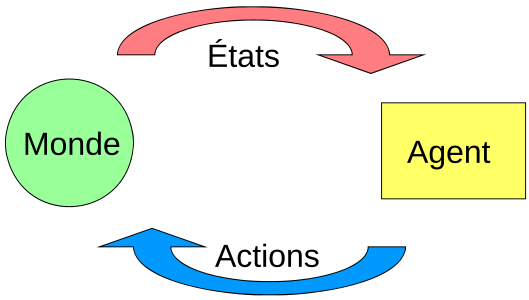
Université Paris 6

Patrice.Perny@lip6.fr

1. Processus Décisionnels Markoviens

Processus Décisionnels Markoviens

Représentation d'une interaction synchrone
entre un agent et le monde

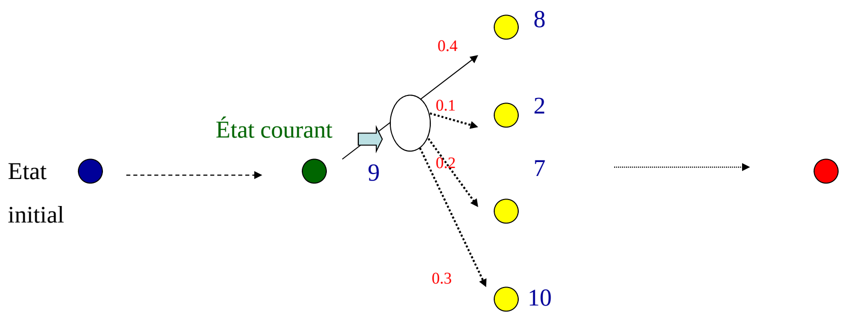


Planification des actions d'un agent dans l'incertain

Multiples Applications

Robotique : sélection de l'action, planification
Gestion de Stocks : planification des commandes
Jeux : calcul de stratégies
…

Représentation dans un graphe d'état



Formalisation d'un PDM

$$\boxed{PDM = \langle S, A, T, R \rangle}$$

S : un ensemble fini d'états du monde

A : un ensemble fini d'actions

$T : S \times A \to L(S)$ fonction de transition

$(s, a) \mapsto T(s, a)$ loi de proba sur les états de la nature

Abus de notation : $T(s, a, s') = T(s, a)(s'), \quad s' \in S$

$R : S \times A \to \mathbb{R}$

$(s, a) \mapsto R(s, a)$ récompense immédiate

Décisions et politiques

Règles de décision : « si l'état est *s* alors exécuter l'action ***a*** »

Représentation par une fonction de décision

$$d : S \to A$$
$$s \mapsto a = d(s) \text{ action choisie dans l'état } s$$

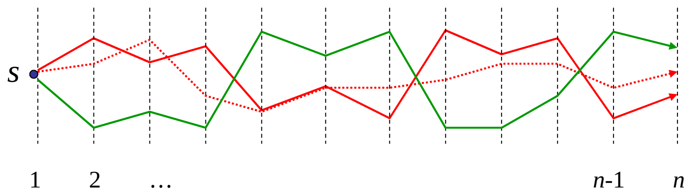
A l'instant $n+1-t$: états possibles S_t, actions possibles A_t

$$d_t : S_t \to A_t$$
$$s \mapsto a = d_t(s) \text{ action choisie à l'instant } n+1-t$$

Hypothèse : *observabilité totale* (on connaît l'état courant)

Décision dynamique

n étapes de décision (n = horizon, fini ou infini)



politique = séquence de fonctions de décisions $(d_n, \ldots, d_1)$

stratégie stationnaire : $(d, d, \ldots, d)$

Critères à optimiser

Horizon fini :

$$E\left(\sum_{t=1}^{N} R(s, d_t(s))\right)$$

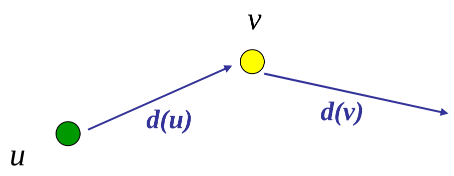
Horizon infini :

$$E\left(\sum_{t=1}^{N} \gamma^t R(s, d_t(s))\right)$$

2. Résolution de PDMs

ÉVALUATION LOCALE D'UNE POLITIQUE

Le cas déterministe

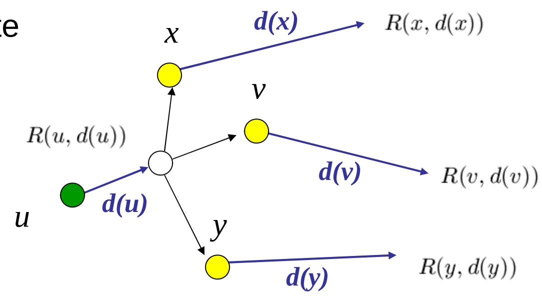


à horizon 1 à horizon 2

récompense : $R(u, d(u)) \;\; + R(v, d(v))$

ÉVALUATION LOCALE D'UNE POLITIQUE

Le cas non-déterministe



à horizon 1

$V_d(u) = R(u, d(u))$

à horizon 2

$$V_d(u) = R(u, d(u))$$
$$+\gamma \;\; [T(u, d(u), x)R(x, d(x)) + T(u, d(u), v)R(v, d(v)) + T(u, d(u), y)R(y, d(y))]$$

facteur de réduction : $0 < \gamma \leq 1$ $\quad$ ($\gamma < 1$ si horizon infini)

Evaluation locale d'une politique

Dernière décision

$$V_{d,1}(s) = R(s, d_1(s))$$

Décision à t étapes de la fin

$$V_{d,t}(s) = R(s, d_t(s)) + \gamma \sum_{s' \in S} T(s, d_t(s), s') V_{d,t-1}(s')$$

Calcul des décisions optimales (horizon fini)

Dernière décision

$$V_1^*(s) = \max_a R(s,a) \qquad d_1^*(s) = \arg\max_a R(s,a)$$

Décision à t étapes de la fin

$$V_t^*(s) = \max_a \left[R(s,a) + \gamma \sum_{s' \in S} T(s,a,s') V_{t-1}^*(s') \right]$$

$\gamma = 1$

$$d_t^*(s) = \arg\max_a \left[R(s,a) + \gamma \sum_{s' \in S} T(s,a,s') V_{t-1}^*(s') \right]$$

Algorithme 1 : induction arrière

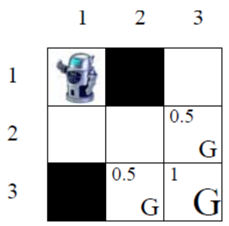
Résolution d'un PDM à horizon fini N

Algorithm 1: Backward Induction

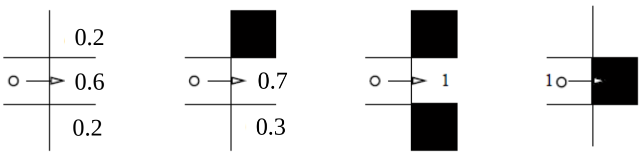
foreach $s \in S$ **do** $V_0^*(s) = 0$
for $t \leftarrow 1$ **to** N **do**
  for $s \in S$ **do**
    for $a \in A$ **do**
      $Q_t^a(s) = R(s,a) + \sum_{s' \in S} T(s,a,s') V_{t-1}^*(s')$
    end
    $a^* \leftarrow \text{choice}[\text{argmax}_{a \in A}\, Q_t^a(s)]$
    $d_t(s) \leftarrow a^*$
    $V_t^*(s) \leftarrow Q_t^{a^*}(s)$
  end
end
return $d_t(s), V_t^*(s)$ for all $t = 1, \ldots, N, s \in S$

choice (X) : choix d'un élément dans X

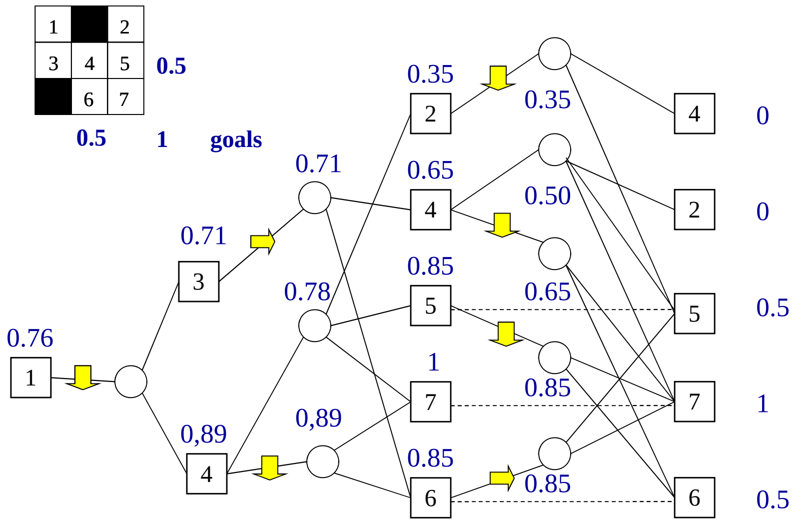
Exemple



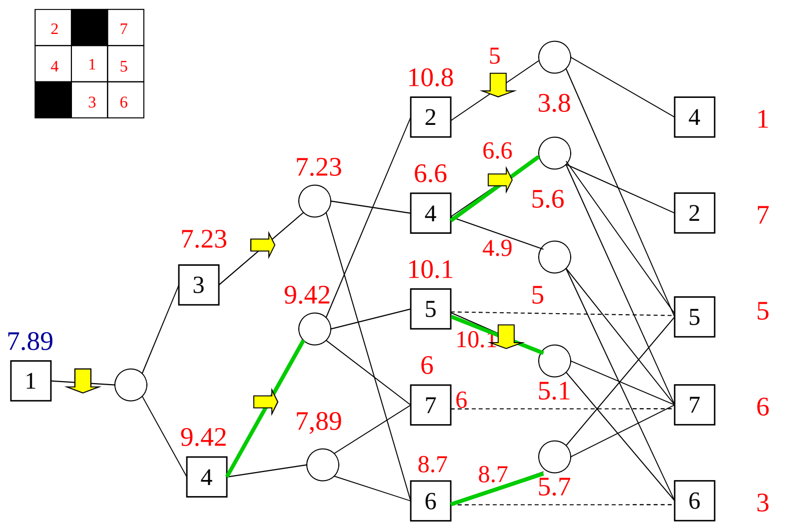
Possibilités de transition



CALCUL DE EU DANS LE GRAPHE DE DÉCISION



CALCUL DE EU DANS LE GRAPHE DE DÉCISION



Résolution d'un PDM à horizon infini

Evaluation d'une stratégie stationnaire : $(d, d, \ldots, d)$

$$V_d(s) = R(s, d(s)) + \gamma \sum_{s' \in S} T(s, d(s), s') V_d(s'), \quad s \in S$$

Théorème (Howard, 1960)

Il existe une stratégie stationnaire optimale pour tout état initial.

La valeur de cette stratégie est caractérisée par le système:

$$V^*(s) = \max_a \left[R(s, a) + \gamma \sum_{s' \in S} T(s, a, s') V^*(s') \right], \quad s \in S$$

qui admet une solution unique.

Equations de Bellman

Formulation Vectorielle et approche itérative

$$\forall s \in S, LV(s) = \max_{a \in A} \left\{ R(s, a) + \gamma \sum_{s' \in S} T(s, a, s') V(s') \right\}$$

$$LV = \max_\pi \{R_\pi + \gamma T_\pi V\}$$

Equations de Bellman : $V = LV$

Proposition : $\|LV - LV'\| \leq \gamma \|V - V'\|$

Théorème du point fixe de Banach $\Rightarrow$
la suite $\{V_n\}$ définie par :$V_0 = 0, \quad V_t = LV_{t-1}$ for all $t \geq 1$
converge vers la solution V^* de l'équation $V = LV$

$\rightarrow$ calculer $V_0, V_1, \ldots, V_t$ jusqu'à $\|V_t - V_{t-1}\| < \varepsilon$

Algorithme 2 : itération de la valeur

Algorithme de l'itération de la valeur (Bellman, 57)

Algorithm 2: Value Iteration

foreach $s \in S$ **do** $V_0(s) = 0$
$t \leftarrow 0$
repeat
  $t \leftarrow t+1$
  for $s \in S$ **do**
    for $a \in A$ **do**
      $Q_t^a(s) = R(s,a) + \gamma \sum_{s' \in S} T(s,a,s') V_{t-1}(s')$
    end
    $V_t(s) \leftarrow \max_{a \in A} Q_t^a(s)$
  end
until $\max_{s \in S}\{|V_t(s) - V_{t-1}(s)|\} < \varepsilon$;
foreach $s \in S$ **do** $d(s) \leftarrow \text{choice}[\text{argmax}_{a \in A} Q_t^a(s)]$
return $d(s), V_t(s)$ for all $s \in S$

Amélioration : Gauss-Seidel (utiliser V_t(s') à la place de V_{t-1}(s') si déjà calculé

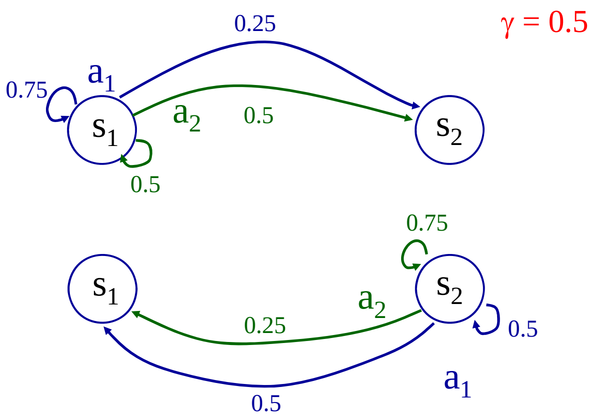
Garantie de performance

Borne de l'erreur (Williams and Baird, 1993)

Si $|V_t(s) - V_{t-1}(s)| < \epsilon$ pour tout $s \in S$ alors :

$$\max_{s \in S} |V_{d_{V_t}}(s) - V^*(s)| < 2\epsilon \frac{\gamma}{1-\gamma}$$

Exemple



Rewards	s1	s2
a1	8	11
a2	12	9

Itération de la valeur

$$
\begin{aligned}
v_1 &= \max\{8 + \tfrac{3}{8}v_1 + \tfrac{1}{8}v_2, 12 + \tfrac{1}{4}v_1 + \tfrac{1}{4}v_2\} \\
v_2 &= \max\{11 + \tfrac{1}{4}v_1 + \tfrac{1}{4}v_2, 9 + \tfrac{1}{8}v_1 + \tfrac{3}{8}v_2\}
\end{aligned}
$$

v1	v2	Q11	Q12	Q21	Q22
0	0				
12,00	11,00	8,00	12,00	11,00	9,00
17,75	16,75	13,88	17,75	16,75	14,63
20,63	19,63	16,75	20,63	19,63	17,50
22,06	21,06	18,19	22,06	21,06	18,94
22,78	21,78	18,91	22,78	21,78	19,66
23,14	22,14	19,27	23,14	22,14	20,02
23,32	22,32	19,45	23,32	22,32	20,20
23,41	22,41	19,54	23,41	22,41	20,29
23,46	22,46	19,58	23,46	22,46	20,33
23,48	22,48	19,60	23,48	22,48	20,35
23,49	22,49	19,61	23,49	22,49	20,36
23,49	22,49	19,62	23,49	22,49	20,37
23,50	22,50	19,62	23,50	22,50	20,37
23,50	22,50	19,62	23,50	22,50	20,37

L'EXEMPLE DU ROBOT AVEC HORIZON INFINI

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Probas							Bas				Droite				Altitude	
2																	
3	0,2	1	0,00		0,00		-1,00		-1,00		0,00		0,00		1		3
4	0,6		0,00	0,00	0,00		0,00	33,00	14,00		-1,00	9,00	0,00		2	3	4
5	0,2			0,00	100,00			0,00	100,00			15,00	100,00			5	3
6																	
7		2	0,00		0,00		10,55		12,65		0,00		0,00				
8	0,3		0,00	33,00	14,00		0,00	35,25	19,25		9,20	13,20	7,00				
9	0,7			15,00	100,00			7,50	100,00			19,90	100,00				
10																	
11	Gamma	3	10,55		12,65		12,72		14,23		5,28		6,33				
12	0,5		9,20	35,25	19,25		4,60	35,99	20,97		11,25	16,04	9,63				
13				19,90	100,00			9,95	100,00			21,74	100,00				
14																	
15		4	12,72		14,23		13,28		14,74		6,36		7,11				
16			11,25	35,99	20,97		5,63	36,26	21,61		12,01	16,71	10,48				
17				21,74	100,00			10,87	100,00			22,34	100,00		Politique optimale		
18																	
19		5	13,28		14,74		13,49		14,93		6,64		7,37		v		v
20			12,01	36,26	21,61		6,00	36,35	21,82		12,26	16,96	10,80		>	v	v
21				22,34	100,00			11,17	100,00			22,56	100,00			>	>

L'EXEMPLE DU ROBOT AVEC HORIZON INFINI

5	13,28		14,74	13,49		14,93	6,64		7,37	v		v
	12,01	36,26	21,61	6,00	36,35	21,82	12,26	16,96	10,80	>	v	v
		22,34	100,00		11,17	100,00		22,56	100,00		>	>
6	13,49		14,93	13,56		15,00	6,75		7,47			
	12,26	36,35	21,82	6,13	36,38	21,90	12,35	17,04	10,91			
		22,56	100,00		11,28	100,00		22,64	100,00			
7	13,56		15,00	13,59		15,02	6,78		7,50			
	12,35	36,38	21,90	6,17	36,40	21,92	12,38	17,07	10,95			
		22,64	100,00		11,32	100,00		22,66	100,00			
8	13,59		15,02	13,60		15,03	6,79		7,51			
	12,38	36,40	21,92	6,19	36,40	21,93	12,39	17,08	10,96			
		22,66	100,00		11,33	100,00		22,67	100,00			

Autre approche : itération de la politique

Proposition : soit π une politique stationnaire et V_γ^π sa valuation en chaque état. Alors toute politique π' choisie dans $\text{argmax}_\delta\{R_\delta + \gamma T_\delta V_\gamma^\pi\}$ vérifie l'inégalité $V_\gamma^{\pi'} \geq V_\gamma^\pi$

$\Rightarrow$ idée d'un autre algo itératif

Partir avec une politique arbitraire π_0

résoudre le système $V_0 = L_{\pi_0} V_0$

choisir π_1 dans $\text{argmax}_\delta\{R_\delta + \gamma T_\delta V_0\}$

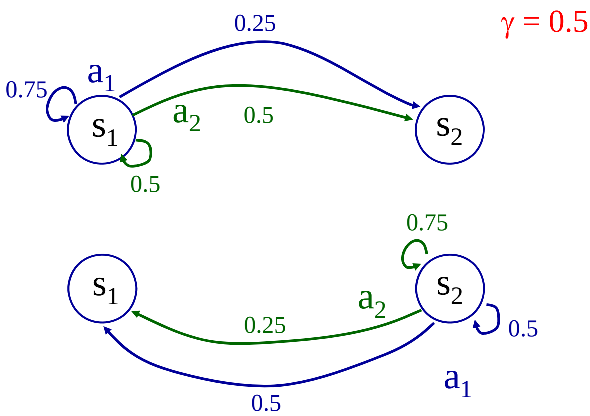
etc... $\quad V_0 \leq V_1 \leq \ldots \leq V_t$

Algorithme 3 : itération de la politique

Algorithm 3: Policy Iteration

choose an initial decision rule d_0
$t \leftarrow 0$
repeat
  résoudre
  $\{V_t(s) = R(s, d_t(s)) + \gamma \sum_{s' \in S} T(s, d_t(s), s') V_t(s'), \quad \forall s \in S\}$
  for $s \in S$ **do**
    $d_{t+1}(s) \leftarrow$choice[$\text{argmax}_{a \in A}\{R(s,a) + \gamma \sum_{s' \in S} T(s, a, s') V_t(s')\}$]
  end
  $t \leftarrow t + 1$
until $d_t(s) = d_{t+1}(s), \forall s \in S$;
return $d_{t+1}(s), V_t(s)$ for all $s \in S$

Exemple



Rewards	s1	s2
a1	8	11
a2	12	9

Exemple

Evaluation de la politique : a_1 si s_1 et a_2 si s_2

Résoudre le système :

$$\begin{aligned} v_1 &= 8 + \tfrac{3}{8} v_1 + \tfrac{1}{8} v_2 \\ v_2 &= 9 + \tfrac{1}{8} v_1 + \tfrac{3}{8} v_2 \end{aligned}$$

Utiliser la solution pour mettre à jour la politique courante

Itérer jusqu'à stabilité…

Formulation par PL

Proposition : $V \geq LV \Rightarrow V \geq V^*$

Donc on peut résoudre le problème d'optimisation

$\min \sum_{s\in S} V(s)$

s.c. $V \geq LV$

Résolution par PL

Algorithm 4: Linear Programming Formulation

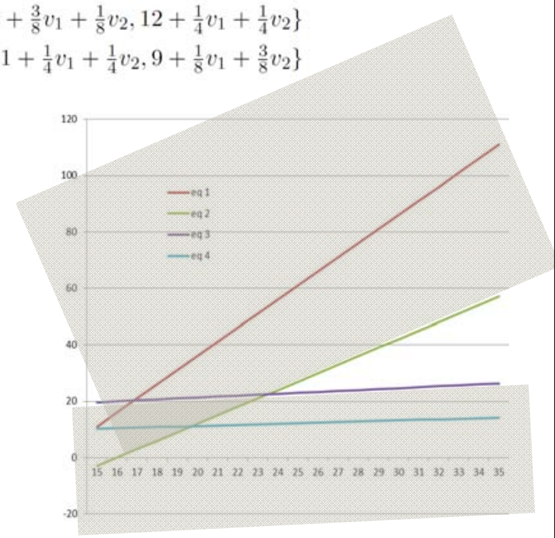
résoudre $\min_V \sum_{s\in S} V(s)$
s.c.
$V(s) \geq R(s,a) + \gamma \sum_{s'\in S} T(s,a,s')V^*(s'), \quad \forall s \in S, \forall a \in A$
foreach $s \in S$ **do**
  $d(s) \leftarrow \text{choice}[\text{argmax}_{a\in A}\{R(s,a) + \gamma \sum_{s'\in S} T(s,a,s')V(s')\}]$
end
return $d(s), V(s)$ for all $s \in S$

Exemple

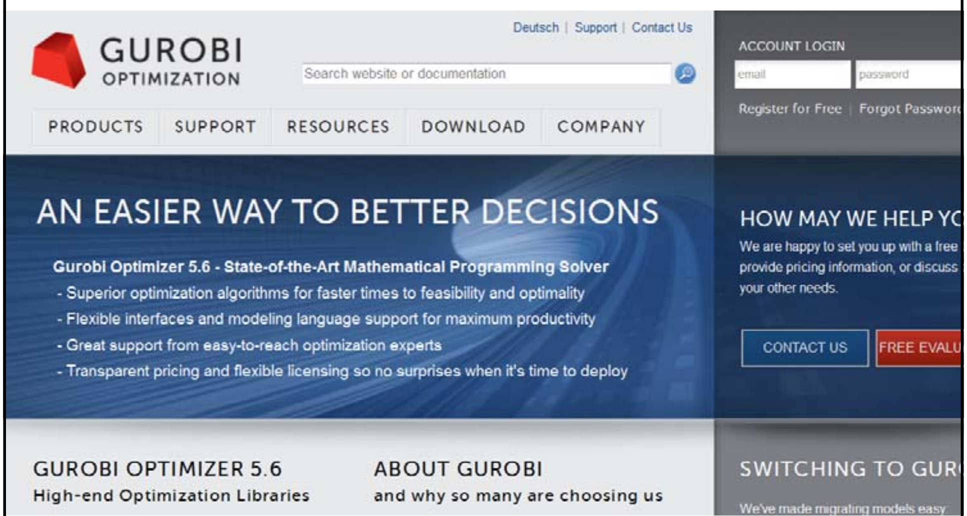
$$\begin{aligned} v_1 &= \max\{8 + \tfrac{3}{8}v_1 + \tfrac{1}{8}v_2, 12 + \tfrac{1}{4}v_1 + \tfrac{1}{4}v_2\} \\ v_2 &= \max\{11 + \tfrac{1}{4}v_1 + \tfrac{1}{4}v_2, 9 + \tfrac{1}{8}v_1 + \tfrac{3}{8}v_2\} \end{aligned}$$

$$\min v_1 + v_2$$

$$s.c. \begin{cases} v_1 \geq 8 + \frac{3}{8}v_1 + \frac{1}{8}v_2 & (1) \\ v_1 \geq 12 + \frac{1}{4}v_1 + \frac{1}{4}v_2 & (2) \\ v_2 \geq 11 + \frac{1}{4}v_1 + \frac{1}{4}v_2 & (3) \\ v_2 \geq 9 + \frac{1}{8}v_1 + \frac{3}{8}v_2 & (4) \end{cases}$$



Résolution par PL



Résolution par PL avec Gurobi

```
Optimize a model with 4 rows, 2 columns and 8 nonzeros
Presolve time: 0.00s
Presolved: 4 rows, 2 columns, 8 nonzeros

Iteration    Objective       Primal Inf.    Dual Inf.      Time
       0    0.0000000e+00   4.000000e+01   0.000000e+00      0s
       2    4.6000000e+01   0.000000e+00   0.000000e+00      0s

Solved in 2 iterations and 0.00 seconds
Optimal objective  4.600000000e+01

Valeurs de la politique optimale:

<gurobi.Var v1 (value 23.5)>
<gurobi.Var v2 (value 22.5)>
Valeur de la fonction objectif : 46.0
```


3. Apprentissage par renforcement dans les PDMs

Le Q-Learning

Cadre des MDPs :

Itération de la valeur, de la politique, PL $\rightarrow$
$v(s)$, $Q(a,s)$, π politique optimale

Imaginons que T ou R n'est pas connu, deux approches :

- approche indirecte (model based) : estimer T et R puis résoudre le MDP
- approche directe (model free) : on va directement essayer d'estimer Q, c'est le Q-learning

Cadre du Q-Learning :

Essayer des actions dans des états, observer, pour apprendre
$Q(a,s)$, $v(s)$, π

Episodes d'observation : $(s, a, r, s', a, r', s'', a'', r'', s''', \ldots)$

- partir avec des estimations initiales $\hat{Q}(a,s)$ des $Q(a,s)$
- mettre à jour les valeurs $\hat{Q}(a,s)$ en fonction des observations faites

Le Q-Learning

Dans le modèle MDP :
$Q(a,s) = R(a,s) + \gamma \sum_{s' \in S} T(s,a,s') \max_{a'} Q(s,a')$

Mais ici on ne connaît plus T et/ou R

Imaginons que l'on fasse l'observation (s, a, s', r)
$\rightarrow$ nouvelle estimation de $\hat{Q}(s,a)$: $r + \gamma \max_{a'} \hat{Q}(s',a')$

Incorporation de cette nouvelle estimation :
$\hat{Q}(s,a) \leftarrow (1-\alpha_t)\hat{Q}(s,a) + \alpha_t[r + \gamma \max_{a'} \hat{Q}(s',a')]$

avec $\alpha_t = 1/t$ où une fonction de ce type qui décroît avec t
(α_t : taux d'apprentissage)

$\hat{Q}(s,a) \leftarrow \hat{Q}(s,a) + \alpha_t[r + \gamma \max_{a'} \hat{Q}(s',a') - \hat{Q}(s,a)]$

L'algorithme de Q-Learning

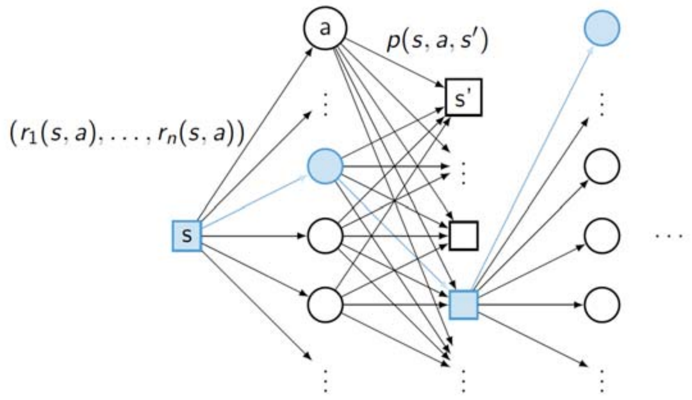
(Watkins, 1989)

Initialize $Q(s, a)$ arbitrarily
Repeat (for each episode):
  Initialize s
  Repeat (for each step of episode):
    Choose a from s using policy derived from Q (e.g., ε-greedy)
    Take action a, observe r, s'
    $Q(s, a) \leftarrow Q(s, a) + \alpha\left[r + \gamma \max_{a'} Q(s', a') - Q(s, a)\right]$
    $s \leftarrow s'$;
  until s is terminal

4. Processus décisionnels Markoviens multiobjectifs

Processus Markoviens Multiobjectifs

Minimiser distance, temps, consommation d'énergie, risque encouru…



Formulations PL pour un seul objectif

Programme primal

$$(\mathcal{P}) \begin{cases} \min \sum_{s \in S} \mu(s) v(s) \\ \text{s.t. } v(s) - \gamma \sum_{s' \in S} T(s, a, s') v(s') \geq R(s, a) \qquad \forall s \in S, \forall a \in A \end{cases}$$

Programme dual

$$(\mathcal{D}) \begin{cases} \max \quad \sum_{s \in S} \sum_{a \in A} R(s, a)\, x_{sa} \\ \text{s.t. } \sum_{a \in A} x_{sa} - \gamma \sum_{s' \in S} \sum_{a \in A} T(s', a, s)\, x_{s'a} = \mu(s) \qquad \forall s \in S \\ x_{sa} \geq 0 \quad \forall s \in S, \forall a \in A \end{cases}$$

Charactérisaton des politiques mixtes

$$\text{s.t.} \sum_{a \in A} x_{sa} - \gamma \sum_{s' \in S} \sum_{a \in A} T(s', a, s)\, x_{s'a} = \mu(s) \qquad \forall s \in S$$
$$x_{sa} \geq 0 \qquad \forall s \in S, \forall a \in A$$

Proposition 1 *For a policy π, if x^π is defined as $x^\pi(s,a) = \sum_{t=0}^{\infty} \gamma^t p_t^\pi(s,a), \forall s \in S, \forall a \in A$ where $p_t^\pi(s,a)$ is the probability of reaching state s and choosing a at step t, then x^π is a feasible solution of $\mathcal{D}$.*

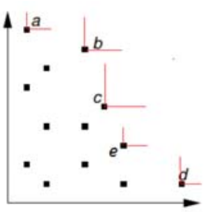
Proposition 2 *If x_{sa} is a solution of $\mathcal{D}$, then the stationary randomized policy δ^∞, defined by $\delta(s,a) = x_{sa} / \sum_{a' \in A} x_{sa'}, \forall s \in S, \forall a \in A$ defines $x^{\delta^\infty}(s,a)$ as in Proposition 1, that are equal to x_{sa}.*

Les solutions du polyèdre des contraintes caractérisées par les variables x_{sa} sont les politiques mixtes du MDP

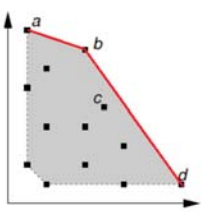
PL et Dualité pour les MOMDPs

$$\max f_i(x) = \sum_{s \in S} \sum_{a \in A} R_i(s,a)\, x_{sa} \qquad \forall i = 1, \ldots, n$$

$$\text{s.t.} \sum_{a \in A} x_{sa} - \gamma \sum_{s' \in S} \sum_{a \in A} T(s', a, s)\, x_{s'a} = \mu(s) \qquad \forall s \in S$$
$$x_{sa} \geq 0 \qquad \forall s \in S, \forall a \in A$$



politiques pures



politiques mixtes

Pour forcer des politiques pures

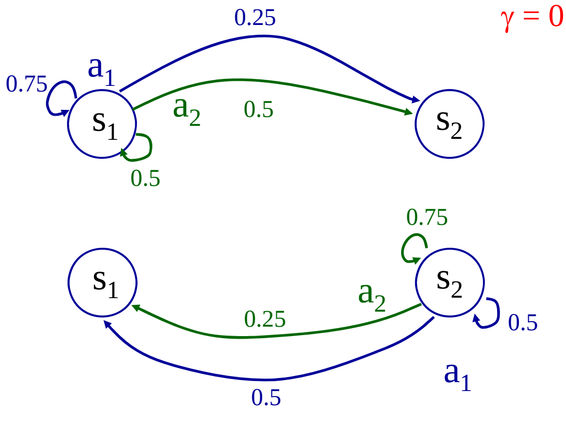
$$\sum_{a \in A} d_{sa} \leq 1 \qquad \forall s \in S$$
$$(1-\gamma)x_{sa} \leq d_{sa} \qquad \forall s \in S, \forall a \in A$$
$$d_{sa} \in \{0,1\} \qquad \forall s \in S, \forall a \in A$$

$$x_{sa} = \sum_{t=0}^{\infty} \gamma^t p_t^\pi(s,a) \leq \frac{1}{1-\gamma}$$

Politique Max-min optimale

$$
\begin{aligned}
&\max z \\
&z \leq f_i(x), i = 1, \dots, n \\
&f_i(x) = \sum_{s \in S} \sum_{a \in A} R_i(s,a)\, x_{sa} \quad \forall i = 1, \dots, n \\
&\text{s.t.} \sum_{a \in A} x_{sa} - \gamma \sum_{s' \in S} \sum_{a \in A} T(s', a, s)\, x_{s'a} = \mu(s) \quad \forall s \in S \\
&\qquad x_{sa} \geq 0 \quad \forall s \in S, \forall a \in A
\end{aligned}
$$

Exemple



$\gamma = 0.5$

R1	s1	s2
a1	8	11
a2	12	9

R2	s1	s2
a1	13	7
a2	6	15

LES CONTRAINTES CHANGENT ENTRE R1 ET R2

$$\min \frac{1}{2}v_1 + \frac{1}{2}v_2$$
$$\left\{\begin{array}{rrcr} \frac{5}{8}v_1 & -\frac{1}{8}v_2 & \geq & 8 \\ \frac{3}{4}v_1 & -\frac{1}{4}v_2 & \geq & 12 \\ -\frac{1}{4}v_1 & +\frac{3}{4}v_2 & \geq & 11 \\ -\frac{1}{8}v_1 & +\frac{5}{8}v_2 & \geq & 9 \end{array}\right.$$

$$\min \frac{1}{2}v_1 + \frac{1}{2}v_2$$
$$\left\{\begin{array}{rrcr} \frac{5}{8}v_1 & -\frac{1}{8}v_2 & \geq & 13 \\ \frac{3}{4}v_1 & -\frac{1}{4}v_2 & \geq & 6 \\ -\frac{1}{4}v_1 & +\frac{3}{4}v_2 & \geq & 7 \\ -\frac{1}{8}v_1 & +\frac{5}{8}v_2 & \geq & 15 \end{array}\right.$$

FORMES DUALES, LES OBJECTIFS CHANGENT

$$\max 8x_{11} + 12x_{12} + 11x_{21} + 9x_{22}$$
$$\left\{\begin{array}{rrrrcr} \frac{5}{8}x_{11} & +\frac{3}{4}x_{12} & -\frac{1}{4}x_{21} & -\frac{1}{8}x_{22} & = & \frac{1}{2} \\ -\frac{1}{8}x_{11} & -\frac{1}{4}x_{12} & +\frac{3}{4}x_{21} & +\frac{5}{8}x_{22} & = & \frac{1}{2} \end{array}\right.$$
$$x_{11} \geq 0,\ x_{12} \geq 0,\ x_{21} \geq 0,\ x_{22} \geq 0,$$

$$\max 13x_{11} + 6x_{12} + 7x_{21} + 15x_{22}$$
$$\left\{\begin{array}{rrrrcr} \frac{5}{8}x_{11} & +\frac{3}{4}x_{12} & -\frac{1}{4}x_{21} & -\frac{1}{8}x_{22} & = & \frac{1}{2} \\ -\frac{1}{8}x_{11} & -\frac{1}{4}x_{12} & +\frac{3}{4}x_{21} & +\frac{5}{8}x_{22} & = & \frac{1}{2} \end{array}\right.$$
$$x_{11} \geq 0,\ x_{12} \geq 0,\ x_{21} \geq 0,\ x_{22} \geq 0,$$

Optimisation max-min

$$\max z$$

$$\left\{\begin{array}{rrrrrr} z \leq & 8x_{11} & +12x_{12} & +11x_{21} & +9x_{22} & \\ z \leq & 13x_{11} & +6x_{12} & +7x_{21} & +15x_{22} & \\ \frac{5}{8}x_{11} & +\frac{3}{4}x_{12} & -\frac{1}{4}x_{21} & -\frac{1}{8}x_{22} & = & \frac{1}{2} \\ -\frac{1}{8}x_{11} & -\frac{1}{4}x_{12} & +\frac{3}{4}x_{21} & +\frac{5}{8}x_{22} & = & \frac{1}{2} \end{array}\right.$$

$$x_{11} \geq 0,\ x_{12} \geq 0,\ x_{21} \geq 0,\ x_{22} \geq 0,$$

Application :