Examen « Apprentissage orienté agent »

30 janvier 2025 - Durée 2h30 - Documents autorisés

Exercice 1 (11 points)

Dans le cadre les Processus Décisionnels de Markov (PDM), un agent «TOTO» est dans l'un des 3 états de `E = {a, b, c}`. Dans chaque état, il peut effectuer 2 actions appartenant à `A = {1, 2}`.

La fonction de transition `P` de `ExAxE` dans `[0,1]` est définie par:

```
P(a, 1, b) = 1    P(a, 2, b) = P(a, 2, c) = 1/2
P(b, 1, c) = 1    P(b, 2, c) = P(b, 2, a) = 1/2
P(c, 1, b) = 1    P(c, 2, a) = 1   et   P(e, n, e') = 0 sinon.
```

La fonction de récompense `R` de `ExAxE` dans `[0,2]` est définie par:

```
R(b, 2, a) = 1    R(c, 2, a) = 2   et   R(e, n, e') = 0 sinon.
```

Le taux de diminution γ vaut ½.

Q1 : Dessiner le graphe de ce PDM. **(1pt)**

On suppose que l'agent «toto» utilise la *politique aléatoire* `π` sur probabilité uniforme entre les deux actions: `π(e, 1)= π(e, 2) = ½`.

Q2 : Ecrire les équations de Bellman

a) donnant `V` en fonction de `Q`. **(1pt)**

b) donnant `Q` en fonction de `V`. **(1pt)**

Q3 : a) Ecrire les 3 équations linéaires ayant V(a), V(b) et V(c) pour inconnues. **(0.5 pt)**

b) Eliminer V(c) pour obtenir 2 équations ayant V(a) et V(b) pour inconnues. **(1 pt)**

c) En déduire V(a) et V(b) puis V(c). **(1 pt)**

d) Donner les valeurs de la fonction `Q`. **(1 pt)**

On suppose que l'agent «toto» utilise la *politique optimale* `π*`.

Q4 : a) Ecrire les équations de Bellman donnant `Q*` en fonction de `V*`. **(1pt)**

b) Donner des arguments expliquant pourquoi, pour chaque état `e`, il est raisonnable de faire l'hypothèse que `Q*(e, 1) < Q*(e, 2)`. **(1 pt)**

c) Avec ces hypothèses, écrire les équations donnant `V*` en fonction de `Q*`. **(0.5 pt)**

Q5 : a) Résoudre le système en `V*`. **(1.5 pt)**

b) Donner les valeurs de la fonction `Q*` et vérifier les hypothèses faites en 4b). **(1pt)**

Exercice 2 (5 points)

Soit **B** le jeu matriciel de la table 1 :

	c	d
a	1, 1	100, 2
b	2, 100	3, 3

Table 1

1° a) Quels sont les équilibres de Nash purs ? **(1 pt)**

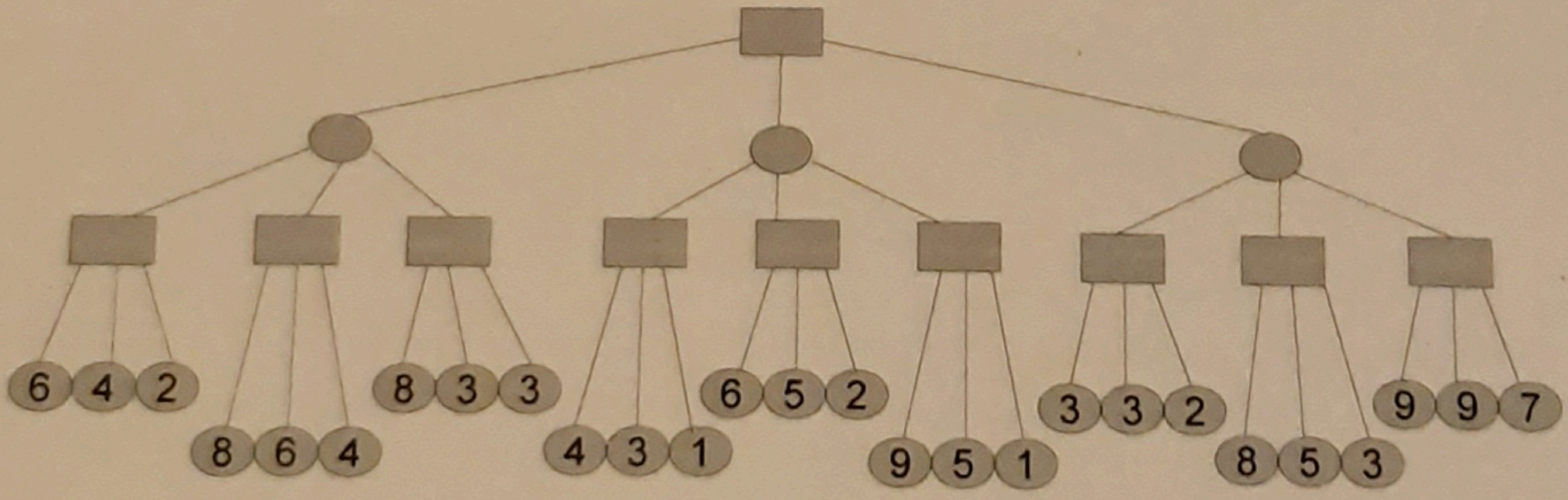
b) le(s) équilibre(s) mixte(s) ? Quels sont les retours à l'équilibre mixte ? **(2 pts)**

2° a) Pourquoi D(a,d)=D(b, c)=½ et D(a, c)=D(b, d)=0 est-il un équilibre corrélé ? **(1 pt)**

b) Que valent les retours à l'équilibre ? **(1 pt)**

Exercice 3 (4 points)

Soit A l'arbre d'un jeu à 2 joueurs, Max et Min. Max est représenté par des rectangles et Min par des ovales.



Sur la feuille de réponse :

Pour Minimax, donner les valeurs minimax de chaque nœud.

Pour alfa-béta [-∞, +∞], alfa-béta [5, 6], et alfa-béta [6, 7], les nœuds sont explorés de gauche à droite.

Donner les les coupes alfa-béta et les valeurs alfa et béta de chaque nœud exploré. **(1 pt par cas)**