

## Série 8

Exercice 1:

- Car driving: Markov Decision Process, parce que l'état suivant dépend seulement de l'état actuel et de la décision prise (tourner, freiner, etc)
- Playing poker: Non-Markovian Decision Process, parce que l'état suivant dépend de l'historique de la partie (pas seulement le coup d'avant) et de la décision prise (se coucher, relancer, etc)
- Stock: Non-Markovian Process, parce que cela dépend du prix actuel mais aussi de pleins de facteurs différents (économie, tendance, etc)

Exercice 2:

1.  $E[U|a=x] = 0 + (0,3 \cdot 0) + (0,7 \cdot 1) = 0,7$
2.  $E[U|a=y] = 1 + (0,6 \cdot 0) + (0,4 \cdot 1) = 1,4$
3.  $V_t^*(s_t) = (0,7 \cdot 0,2) + (1,4 \cdot 0,8) = 1,26$
4.  $E[U|a=y] > E[U|a=x]$  donc y est optimal

Exercice 3:

1. Pour les actions  $a_1$ :

$$\begin{array}{lll} P(s_1 \rightarrow s_1 | a_1) = 0 & P(s_2 \rightarrow s_1 | a_1) = 0 & P(s_3 \rightarrow s_1 | a_1) = 0 \\ P(s_1 \rightarrow s_2 | a_1) = 0,8 & P(s_2 \rightarrow s_2 | a_1) = 1 & P(s_3 \rightarrow s_2 | a_1) = 0 \\ P(s_1 \rightarrow s_3 | a_1) = 0,2 & P(s_2 \rightarrow s_3 | a_1) = 0 & P(s_3 \rightarrow s_3 | a_1) = 1 \end{array}$$

Pour les actions  $a_2$ :

$$\begin{array}{lll} P(s_1 \rightarrow s_1 | a_2) = 0 & P(s_2 \rightarrow s_1 | a_2) = 0,5 & P(s_3 \rightarrow s_1 | a_2) = 1 \\ P(s_1 \rightarrow s_2 | a_2) = 1 & P(s_2 \rightarrow s_2 | a_2) = 0 & P(s_3 \rightarrow s_2 | a_2) = 0 \\ P(s_1 \rightarrow s_3 | a_2) = 0 & P(s_2 \rightarrow s_3 | a_2) = 0,5 & P(s_3 \rightarrow s_3 | a_2) = 0 \end{array}$$

2. Pour les actions  $a_1$ :

$$r(s_1, a_1) = 0 \quad r(s_2, a_1) = 0 \quad r(s_3, a_1) = 1$$

Pour les actions  $a_2$ :

$$r(s_1, a_2) = 0 \quad r(s_2, a_2) = 2 \quad r(s_3, a_2) = 0$$

3. 

t	$V_t^*(s_1)$	$V_t^*(s_2)$	$V_t^*(s_3)$	$Q_t^*(s_1, a_1)$	$Q_t^*(s_1, a_2)$	$Q_t^*(s_2, a_1)$	$Q_t^*(s_2, a_2)$	$Q_t^*(s_3, a_1)$	$Q_t^*(s_3, a_2)$
3	0	2	1	0	0	0	2	1	0
2	2	2.5	2	1.8	2	2	2.5	2	0
1	2.5	4	3	2.4	2.5	2.5	4	3	2

Calcul:

- $Q_2^*(s_1, a_1) = r(s_1, a_1) + (P(s_2|s_1, a_1) \cdot V_3^*(s_2)) + (P(s_3|s_1, a_1) \cdot V_3^*(s_3)) = 0 + (0,8 \cdot 2) + (0,2 \cdot 1) = 1,8$
- $Q_2^*(s_1, a_2) = r(s_1, a_2) + (P(s_2|s_1, a_2) \cdot V_3^*(s_2)) = 2$
- $Q_2^*(s_2, a_1) = r(s_2, a_1) + (P(s_2|s_2, a_1) \cdot V_3^*(s_2)) = 2$
- $Q_2^*(s_3, a_1) = r(s_3, a_1) + (P(s_3|s_3, a_1) \cdot V_3^*(s_3)) = 2$
- $Q_2^*(s_3, a_2) = r(s_3, a_2) + (P(s_1|s_3, a_2) \cdot V_3^*(s_1)) = 0$

Fait avec Chatgpt car les calculs sont similaires à ceux d'avant (appliquer la formule)

4. ●