23/01/2024 14:21 RL class - quiz on chapter 2

## RL class - quiz on chapter 2 Total des points 3/7 ?



Student last name *	
Igbida	
Student first name *	
Rayanne	

$V^*(s) = \max_{\pi} \mathbb{E}[\sum_t \gamma^t R_t   \pi], \ \forall s \in S$	${}^{\flat}Q(s,a) = \mathbb{E}_{r,s'}[r + \max_{a}' Q(s',a')],$ $\forall s, a \in S \times A$
Option 1	Option 2
$\dot{V}(s) = \max_{a} [r(s, a) + \gamma \sum_{s'} p(s' s, a)V(s')], \forall s \in S$	${}^{\flat}Q(s,a) = r(s,a) +$ $\gamma \int_{S \times A} Q(s',a') \pi(a' s') p(s' s,a) ds' da',$ $\forall s,a \in S \times A$

X Question 2 \*0/1

Value iteration computes a sequence of value functions. Pick the answer(s) that correctly describe this sequence (notations are those used in class).

$$V_{n+1} = T^{\pi}V_n$$

 $Q_{n+1} = T^*Q_n$ 

Option 1

Option 2

Option 4

$$Q_{n+1} = T^{\pi_n} Q_n$$
 with  $\pi_n \in \mathcal{G}Q_n$ 

 $V_{n+1} = T^{\pi_n} V_n$  with  $\pi_n \in \mathcal{G}V_n$ 

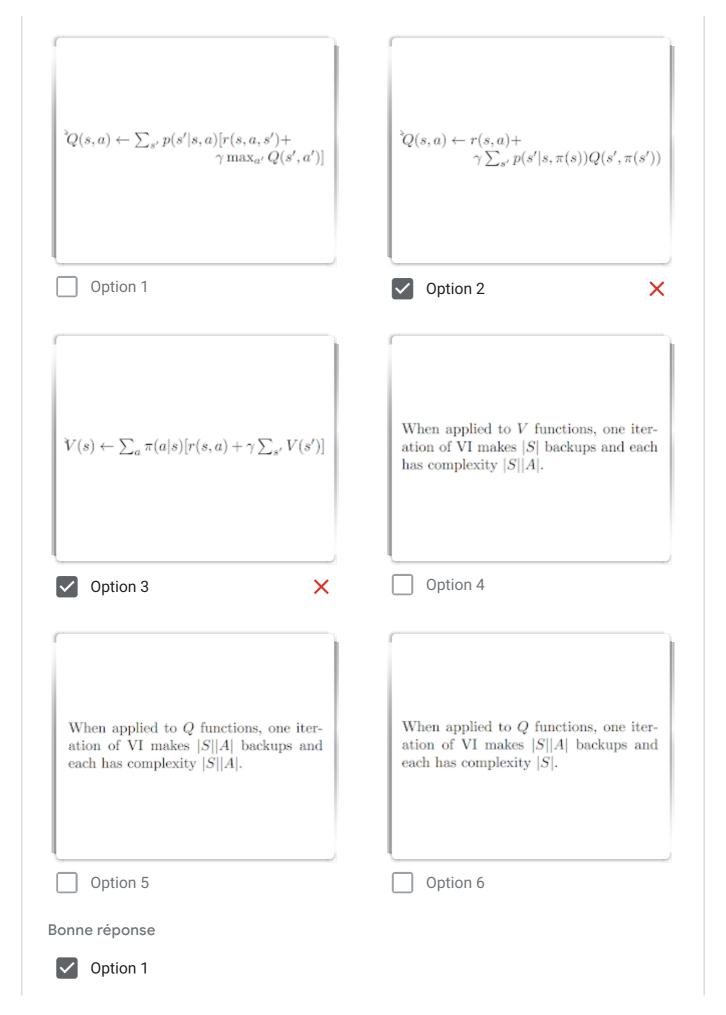
Option 3

Bonne réponse

- Option 2
- Option 3

X Question 3 \*0/1

In MDPs with finite state-action spaces, value iteration (VI) can be written as a sequence of Bellman backups in each state or state-action pair. Pick the correct form(s) and statement(s) for these backups.

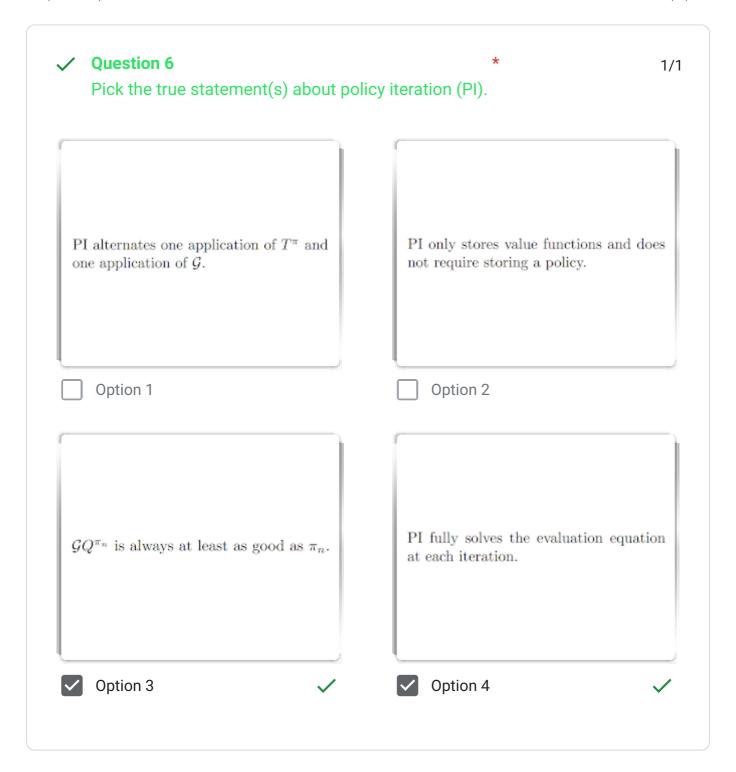


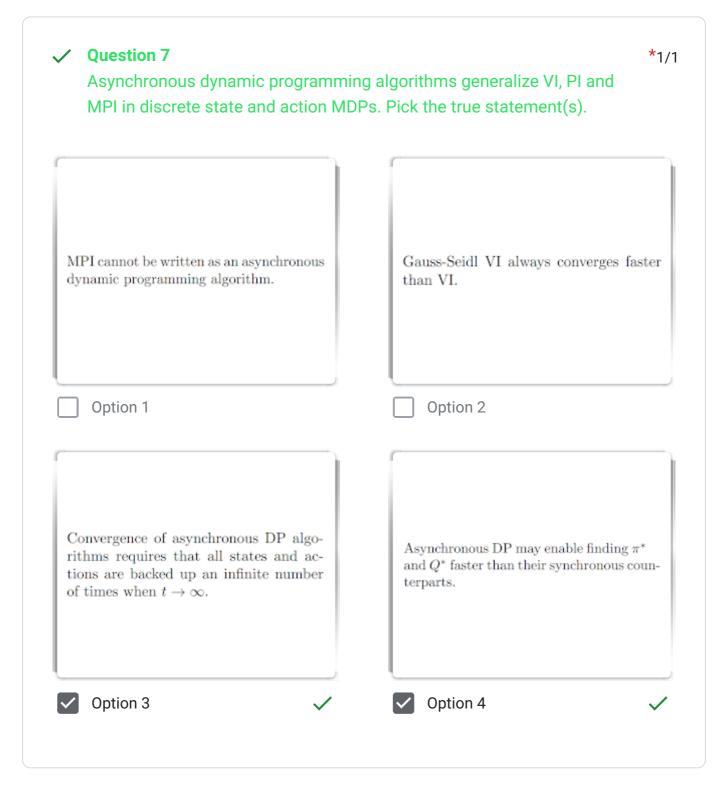
Option 4

Option 6

Question 4 Modified policy iteration (MPI) comp and policies. Pick the true statement MPI.	
$Q_{n+1} = (T^{\pi_n})^m Q_n$ and $\pi_n \in \mathcal{G}Q_n$	Policy iteration is a special case of MPI
Option 1	Option 2
Value iteration is MPI when $m \to \infty$	$Q_{n+1} = (T^*)^m Q_n$ and $\pi_n \in \mathcal{G}Q_n$
Option 3	Option 4
onne réponse	
✓ Option 1	
✓ Option 2	

Any policy that is greedy with respect to $Q^{\pi}$ is optimal.	All policies in $\mathcal{G}Q^*$ are optimal.
Option 1 X	Option 2
There is always only a unique policy in $\mathcal{G}Q^*$ .	$Q^* = \mathcal{G}Q^*$ .
Option 3	Option 4





Ce contenu n'est ni rédigé, ni cautionné par Google. - Conditions d'utilisation - Règles de confidentialité

## Google Forms