Course Discussion Wiki Progress	
Course > Temporal Differenc > Knowledge	Checks > Knowledge Checks
Knowledge Checks □ Bookmark this page	
Question 1	
1/1 point (graded) Which of the following are two characteris	stics of Monte Carlo (MC) and Temporal
Difference (TD) learning?	
TD provides an estimate of after n st	·
whereas TD requires no such model.	ne environment i.e. the transition probabilities,
☑ Both MC and TD are model free metl☐ Both MC and TD use bootstrapping.	hods.
✓	
Submit You have used 1 of 2 attempts	Save Show Answer
✓ Correct (1/1 point)	
Question 2 1/1 point (graded)	
Which of the following are two advantages algorithms?	s of the TD algorithm compared to the MC
☐ No Bias in the estimate of value.	
✓ Works in continuing (non-terminating	g) environments.
☐ Sensitive to initial values.☐ Model free.	
☑ Exploits the Markov Decision Process	s properties to gain efficiency.
Vou bayous od 2 of 2 attornate	
Submit You have used 2 of 2 attempts	Show Answer
✓ Correct (1/1 point)	
Question 3 1/1 point (graded)	
Which of the following statements about stemporal difference (TD) algorithms?	sampling are true for Monte Carlo (MC) and
O Monte Carlo algorithms randomly sa O TD and MC careals states. Stars from	
 TD and MC sample states, St+n, from MC algorithms use deterministic same 	
O TD use random sampling of state-act	
Submit You have used 1 of 2 attempts	Save Show Answer
✓ Correct (1/1 point)	
Question 4 1/1 point (graded)	
Which two of the following describe bias-v	
higher bias.	by sampling until the terminal state, leading to
☑ The MC algorithm reduces bias by sa higher variance.	ampling until the terminal state, leading to
☑ The TD algorithm reduces variance b leading to higher bias.	by sampling a small number of time steps,
☐ The TD algorithm reduces bias by sail to higher variance.	mpling a small number of a time steps, leading
Submit You have used 1 of 2 attempts	
	Save Show Answer
✓ Correct (1/1 point)	
Question 5 1/1 point (graded)	
What is the difference between on-policy a	and off-policy learning?
On-policy learning learns by evaluating	
policy improvement on a target policy by evaluating a target policy and per	ing the results of a behavior policy to perform cy, whereas off-policy learns from experience forming policy improvement on the target
policy improvement on a target policy by evaluating a target policy and perpolicy. On-policy learning learns from experi	rience by evaluating a target policy and
policy improvement on a target policy by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on the second content of t	cy, whereas off-policy learns from experience forming policy improvement on the target
 policy improvement on a target policy by evaluating a target policy and peripolicy. On-policy learning learns from experiperforming policy improvement on the learns by evaluating the results of a known a target policy. On-policy learning learns from experimental contents. 	rience by evaluating a target policy learning the target policy, whereas off-policy, whereas off-policy learning
 policy improvement on a target policy by evaluating a target policy and peripolicy. On-policy learning learns from experiperforming policy improvement on the learns by evaluating the results of a known a target policy. On-policy learning learns from experiperforming policy improvement on the learns of the learns from experiperforming policy improvement on the learns from experiperforming policy improvement on the learns from experiment in the	rience by evaluating a target policy learning behavior policy to perform policy improvement on the target behavior by evaluating a target policy learning behavior policy to perform policy improvement rience by evaluating a target policy and
 policy improvement on a target policy by evaluating a target policy and peripolicy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of a known a target policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on the performing policy improvement on the performance of the performance	rience by evaluating a target policy learning behavior policy to perform policy improvement on the target policy and the target policy to perform policy improvement policy improvement the target policy, whereas off-policy improvement the target policy, whereas off-policy learning the target policy, whereas off-policy learning
 policy improvement on a target policy by evaluating a target policy and peripolicy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of a known a target policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on the performing policy improvement on the performance of the performance	rience by evaluating a target policy learning behavior policy, whereas off-policy improvement on the target behavior policy to perform policy improvement rience by evaluating a target policy improvement behavior policy, whereas off-policy learning the target policy, whereas off-policy learning to perform policy improvement on a behavior policy improvement on a behavior by evaluating a behavior policy and the target policy, whereas off-policy learning the target policy, whereas off-policy learning
 policy improvement on a target policy by evaluating a target policy and peripolicy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of a known a target policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns by evaluating the results of a known in the learns in	rience by evaluating a target policy and the target policy to perform policy learning to perform policy learning behavior policy, whereas off-policy and the target policy, whereas off-policy improvement to perform policy improvement to perform policy learning to perform policy improvement on a behavior policy improvement on a behavior to perform policy improvement on a behavior policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement
 policy improvement on a target policy by evaluating a target policy and peripolicy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of a boundary on a target policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of a boundary on the behavior policy. 	rience by evaluating a target policy and the target policy to perform policy learning behavior policy, whereas off-policy and the target policy, whereas off-policy improvement rience by evaluating a target policy and the target policy, whereas off-policy learning to perform policy improvement on a behavior policy improvement on a behavior rience by evaluating a behavior policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement
 policy improvement on a target policy by evaluating a target policy and peripolicy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of all on a target policy. ✓ On-policy learning learns from experperforming policy improvement on the learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of all on the behavior policy. Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) 	rience by evaluating a target policy and the target policy to perform policy learning behavior policy, whereas off-policy and the target policy, whereas off-policy improvement rience by evaluating a target policy and the target policy, whereas off-policy learning to perform policy improvement on a behavior policy improvement on a behavior rience by evaluating a behavior policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement
policy improvement on a target policy by evaluating a target policy and peripolicy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a lon a target policy. ✓ On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a lon the behavior policy. Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Question 6	rience by evaluating a target policy and the target policy to perform policy improvement on the target policy and the target policy to perform policy improvement rience by evaluating a target policy and the target policy to perform policy improvement rience by evaluating a target policy and the target policy, whereas off-policy learning to perform policy improvement on a behavior rience by evaluating a behavior policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement
policy improvement on a target policy by evaluating a target policy and peripolicy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the results of a learns by evaluating the target policy performing policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the results of a learns by evaluating the target policy policy. Submit You have used 1 of 2 attempts Correct (1/1 point) Puestion 6	rience by evaluating a target policy and che target policy to perform policy improvement on the target policy and che target policy to perform policy improvement rience by evaluating a target policy and che target policy, whereas off-policy learning to perform policy improvement on a behavior perform policy improvement on a behavior rience by evaluating a behavior policy and che target policy, whereas off-policy learning behavior policy to perform policy improvement Save Show Answer
policy improvement on a target polic by evaluating a target policy and peripolicy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of all on a target policy. ✓ On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of all on the behavior policy. Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Question 6 //1 point (graded) Which two statements describe eligibility to Eligibility traces down weight the contents.	rience by evaluating a target policy and the target policy to perform policy improvement on the target policy to perform policy improvement on the target policy to perform policy improvement rience by evaluating a target policy and the target policy, whereas off-policy learning y to perform policy improvement on a behavior policy improvement on a behavior policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement Page 18
policy improvement on a target polic by evaluating a target policy and peripolicy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of a bon a target policy. ✓ On-policy learning learns from experperforming policy improvement on the learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of a bon the behavior policy. Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Question 6 // point (graded) Which two statements describe eligibility the computing average Vs) or Q(s,a).	rience by evaluating a target policy and the target policy to perform policy improvement on the target policy and the target policy to perform policy improvement policy improvement or policy to perform policy improvement or perform policy improvement or perform policy improvement or perform policy improvement or a behavior policy improvement or a behavior policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement or policy improvement or perform policy improvement or
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of all on a target policy. ✓ On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of all on the behavior policy. Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Question 6 //1 point (graded) Which two statements describe eligibility to computing average Vs) or Q(s,a). □ Eligibility traces encourage further expendent of the concomputing average further expendent of the concomputin	rience by evaluating a target policy and the target policy to perform policy improvement on the target policy and the target policy to perform policy improvement policy improvement or policy to perform policy improvement or perform policy improvement or perform policy improvement or perform policy improvement or a behavior policy improvement or a behavior policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement or policy improvement or perform policy improvement or
policy improvement on a target policy by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of all on a target policy. ✓ On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of all on the behavior policy. Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Question 6 //1 point (graded) Which two statements describe eligibility the concomputing average Vs) or Q(s,a). □ Eligibility traces down weight the concomputing average Vs) or Q(s,a). □ Eligibility traces assign credit to action ☑ Eligibility traces assign credit to both states.	rience by evaluating a target policy and the target policy improvement on the target policy and the target policy to perform policy improvement policy improvement on the target policy and the target policy to perform policy improvement on a behavior policy to perform policy improvement on a behavior policy and the target policy and target policy and the target policy and the target
policy improvement on a target policy by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of a bon a target policy. ✓ On-policy learning learns from experperforming policy improvement on the learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of a bon the behavior policy. Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Question 6 1/1 point (graded) Which two statements describe eligibility the concomputing average Vs) or Q(s,a). □ Eligibility traces down weight the concomputing average Vs) or Q(s,a). □ Eligibility traces assign credit to action □ Eligibility traces assign credit to both states. ✓ Submit You have used 1 of 2 attempts	rience by evaluating a target policy and the target policy improvement on the target policy and the target policy to perform policy improvement policy improvement on the target policy and the target policy to perform policy improvement on a behavior policy to perform policy improvement on a behavior policy and the target policy and target policy and the target policy and the target
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a learns by evaluating the results of a learns by evaluating the target policy performing policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a learns by evaluating the target policy. Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) ✓ Correct (1/1 point)	rience by evaluating a target policy learning behavior policy to perform policy improvement on the target policy and the target policy to perform policy improvement on a behavior policy to perform policy improvement on a behavior policy improvement on a behavior policy and the target polic
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the results of a lon a target policy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the results of a lon the behavior policy. Submit You have used 1 of 2 attempts Correct (1/1 point) Question 6 In point (graded) Which two statements describe eligibility to the computing average Vs) or Q(s,a). Eligibility traces down weight the computing average Vs) or Q(s,a). Eligibility traces assign credit to action Eligibility traces assign credit to both states. Submit You have used 1 of 2 attempts Correct (1/1 point) Question 7	rience by evaluating a target policy learning behavior policy to perform policy improvement on the target policy and the target policy to perform policy improvement on a behavior policy to perform policy improvement on a behavior policy improvement on a behavior policy and the target polic
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a bon a target policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a bon the behavior policy. Submit You have used 1 of 2 attempts Correct (1/1 point) Question 6 In point (graded) Which two statements describe eligibility to experiment of the corporation of the policy improvement on the policy. Eligibility traces down weight the corporation of the policy improvement on the policy. Eligibility traces down weight the corporation of the policy improvement on the policy. Eligibility traces assign credit to action of the policy improvement on the	rience by evaluating a target policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement on a behavior policy improvemen
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the results of a learns by evaluating the results of a learns by evaluating the target policy policy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the results of a learns by evaluating the target policy. Submit You have used 1 of 2 attempts Legiphility traces assign credit to both states.	traces? Save Sav
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a lon a target policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a lon the behavior policy. Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Puestion 6 I point (graded) Which two statements describe eligibility to accomputing average Vs) or Q(s,a). □ Eligibility traces down weight the corcomputing average Vs) or Q(s,a). □ Eligibility traces assign credit to action ☑ Eligibility traces assign credit to both states. ✓ Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Question 7 I point (graded) Which of the following characterizes Q-Learning uses bootstrapping and S O Q-Learning uses bootstrapping and S	rience by evaluating a target policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement on the target policy, whereas off-policy learning behavior policy, whereas off-policy learning in the target policy, whereas off-policy learning is to perform policy improvement on a behavior policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement on a behavior policy and the target policy, whereas off-policy learning behavior policy and the target policy and target policy and the targe
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a long a target policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a long the behavior policy. Submit You have used 1 of 2 attempts Vount have used 1 of 2 attempts Eligibility traces down weight the concomputing average Vs) or Q(s,a). Eligibility traces encourage further expended by the concomputing average Vs) or Q(s,a). Eligibility traces assign credit to action Eligibility traces assign credit to action Values to both states. Vount have used 1 of 2 attempts Correct (1/1 point) Question 7 You have used 1 of 2 attempts Vount have used 1 of 2 attempts Submit You have used 1 of 2 attempts Orect (1/1 point) Question 7 You have used 1 of 2 attempts Submit You have used 1 of 2 attempts Submit You have used 1 of 2 attempts Submit You have used 1 of 2 attempts	rience by evaluating a target policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement on the target policy, whereas off-policy learning behavior policy to perform policy improvement rience by evaluating a target policy and the target policy, whereas off-policy learning is to perform policy improvement on a behavior rience by evaluating a behavior policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement behavior policy to perform policy improvement save as a save save save save save save
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a lon a target policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a lon the behavior policy. Submit You have used 1 of 2 attempts Correct (1/1 point) Question 6 If point (graded) Which two statements describe eligibility to computing average Vs) or Q(s,a). Eligibility traces down weight the corcomputing average Vs) or Q(s,a). Eligibility traces assign credit to action Eligibility traces assign credit to both states. Correct (1/1 point) Vou have used 1 of 2 attempts Correct (1/1 point) Question 7 If point (graded) Which of the following characterizes Q-Learning uses bootstrapping and Q-Learning uses bootstrapping and Q-Learning is an on-policy algorithm, Q-Learning is an on-policy algorithm,	rience by evaluating a target policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement on the target policy, whereas off-policy learning behavior policy to perform policy improvement or the target policy, whereas off-policy learning or to perform policy improvement on a behavior policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement or a behavior policy to perform policy improvement or policy to perform policy improvement of the target policy, whereas off-policy learning behavior policy to perform policy improvement or a behavior policy to perform policy improvement or traces? In the most frequently visited and last visited or the most frequently visited and last visited or the most frequently visited and last visited or save arning and SARSA? SARSA does not. SARSA does not. The lagorithms.
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a lon a target policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a lon the behavior policy. Submit You have used 1 of 2 attempts Correct (1/1 point) Question 6 If point (graded) Which two statements describe eligibility to computing average Vs) or Q(s,a). Eligibility traces down weight the concomputing average Vs) or Q(s,a). Eligibility traces assign credit to action Eligibility traces assign credit to both states. Submit You have used 1 of 2 attempts Correct (1/1 point) Question 7 If point (graded) Which of the following characterizes Q-Learning uses bootstrapping and Q-Learning uses bootstrapping and Q-Learning is an on-policy algorithm,	traces? Intribution of states that are rarely visited to save Show Answer save Show Answer The most frequently visited and last visited In the most frequently visited and last visited
policy improvement on a target polic by evaluating a target policy and per policy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the results of all on a target policy. ✓ On-policy learning learns from exper performing policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the results of all on the behavior policy. Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Question 6 1/1 point (graded) Which two statements describe eligibility to computing average Vs) or Q(s,a). □ Eligibility traces down weight the concomputing average Vs) or Q(s,a). □ Eligibility traces assign credit to action ☑ Eligibility traces assign credit to both states. ✓ Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Question 7 1/1 point (graded) Which of the following characterizes Q-Learning uses bootstrapping and S-Calcanning uses bootstrapping and Q-Learning is an on-policy algorithm, when the policy algorithm is an on-policy algorithm, when the policy algorithm is an on-policy algorithm, when the policy algorithm is an on-policy algorithm, when the policy algorithm is an on-policy algorithm, when the policy algorithm is an on-policy algorithm.	traces? tra
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of all on a target policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of all on the behavior policy. Submit You have used 1 of 2 attempts Correct (1/1 point) Question 6 1/1 point (graded) Which two statements describe eligibility the concomputing average Vs) or Q(s,a). Eligibility traces down weight the concomputing average Vs) or Q(s,a). Eligibility traces assign credit to action Eligibility traces assign credit to both states. V Submit You have used 1 of 2 attempts Correct (1/1 point) Question 7 1/1 point (graded) Which of the following characterizes Q-Learning uses bootstrapping and Q-Learning uses bootstrapping and Q-Learning is an on-policy algorithm, when the policy algorithm is an on-policy algorithm, when the correct (1/1 point) Submit You have used 1 of 2 attempts Correct (1/1 point) Submit You have used 1 of 2 attempts Correct (1/1 point)	traces? Intribution of states that are rarely visited to save Show Answer Save Show Answer
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of all on a target policy. ✓ On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of all on the behavior policy. Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Question 6 1/1 point (graded) Which two statements describe eligibility traces down weight the corcomputing average Vs) or Q(s,a). □ Eligibility traces encourage further expendent to action of the policy in the policy algorithm, when the policy in t	rience by evaluating a target policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement on the target policy, whereas off-policy learning behavior policy to perform policy and the target policy, whereas off-policy learning to perform policy improvement on a behavior policy and the target policy, whereas off-policy learning to perform policy improvement on a behavior policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement the target policy to perform policy improvement behavior policy to perform policy improvement behavior policy to perform policy improvement traces? Save Show Answer Save Show Answer Save Show Answer Show Answer Save Show Answer
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of all on a target policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of all on the behavior policy. Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Question 6 1/1 point (graded) Which two statements describe eligibility the corporation of the policy and the corporation of the policy improvement on the learns by evaluating the results of all on the behavior policy. Eligibility traces down weight the corporation of the policy and the corporation of the policy in the policy and the policy and the policy in the policy and the policy and the policy and the policy and the policy algorithm, when the policy and policy algorithm, when the policy algorithm, when the policy and policy algorithm, when the policy and policy algorithm, when the policy algorithm, when the policy and policy algorithm, when the policy and policy algorithm, when the policy and policy algorithm, when the policy algorithm, when the policy and policy algorithm, when the policy and policy algorithm, when the policy and policy.	rience by evaluating a target policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement on the target policy. Whereas off-policy learning behavior policy to perform policy improvement on a behavior policy to perform policy improvement on a behavior policy and the target policy, whereas off-policy learning to perform policy improvement on a behavior policy on a behavior policy improvement on a behavior on a behavior policy and the target policy
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of all on a target policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on the learns by evaluating the results of all on the behavior policy. Submit You have used 1 of 2 attempts Correct (1/1 point) Question 6 1/1 point (graded) Which two statements describe eligibility the concomputing average Vs) or Q(s,a). Eligibility traces down weight the concomputing average Vs) or Q(s,a). Eligibility traces assign credit to action Eligibility traces assign credit to both states. Vou have used 1 of 2 attempts Correct (1/1 point) Question 7 1/1 point (graded) Which of the following characterizes Q-Learning uses bootstrapping and Q-Learning uses bootstrapping and Q-Learning uses bootstrapping and Q-Learning is an on-policy algorithm, when the policy algorithm algorithm and the policy algorithm.	traces? Intribution of states that are rarely visited to sploration of the state space. In the most frequently visited and last visited to sploration of states that are rarely visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state space. In the most frequently visited and last visited to sploration of the state splorat
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a lon a target policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a lon the behavior policy. Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Question 6 1/1 point (graded) Which two statements describe eligibility traces down weight the corcomputing average Vs) or Q(s,a). □ Eligibility traces down weight the corcomputing average Vs) or Q(s,a). □ Eligibility traces assign credit to action ☑ Eligibility traces assign credit to both states. ✓ Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Question 7 1/1 point (graded) Which of the following characterizes Q-Learning uses bootstrapping and Q-Learning uses bootstrapping and Q-Learning is an on-policy algorithm, when the policy algorithm is an on-policy algorithm.	traces? Intribution of states that are rarely visited to save save show Answer The most frequently visited and last visited Interpretation of the state space. Intribution of states that are rarely visited to save show Answer Interpretation of the state space. Intribution of states that are rarely visited to save show Answer Interpretation of the state space. Intribution of states that are rarely visited to save show Answer Interpretation of the state space. Intribution of states that are rarely visited to save show Answer Interpretation of the state space. Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answer Intribution of states that are rarely visited to save show Answe
policy improvement on a target polic by evaluating a target policy and per policy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the results of a lon a target policy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the results of a lon the behavior policy. Submit You have used 1 of 2 attempts Correct (1/1 point) Question 6 1/1 point (graded) Which two statements describe eligibility to be a ligibility traces down weight the corcomputing average Vs) or Q(s.a). Eligibility traces down weight the corcomputing average Vs) or Q(s.a). Eligibility traces assign credit to action Eligibility traces assign credit to both states. V Submit You have used 1 of 2 attempts Correct (1/1 point) Question 7 1/1 point (graded) Which of the following characterizes Q-Learning uses bootstrapping and Q-Learning uses bootstrapping and Q-Learning is an on-policy algorithm, when the policy algorithm is an on-policy algorithm. Submit You have used 1 of 2 attempts Correct (1/1 point)	rience by evaluating a target policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement on the target policy, whereas off-policy learning behavior policy to perform policy improvement or a behavior policy improvement on a behavior policy to perform policy improvement on a behavior policy to perform policy improvement behavior policy to perform policy improvement on the target policy, whereas off-policy learning behavior policy to perform policy improvement on the state space. The most frequently visited and last visited to exploration of the state space. Save Show Answer Save Show Answer All the most frequently visited and last visited to exploration of the state space. The most frequently visited and last visited to exploration of the state space. The most frequently visited and last visited to exploration of the state space. All the most frequently visited and last visited to exploration of the state space. All the most frequently visited and last visited to exploration of the state space. All Rights Reserve save show Answer All Rights Reserve save show Answer
policy improvement on a target polic by evaluating a target policy and per policy. ② On-policy learning learns from exper performing policy improvement on t learns by evaluating the results of a lon a target policy. ✓ ③ On-policy learning learns from exper performing policy improvement on t learns by evaluating the target policy policy. ④ On-policy learning learns from exper performing policy improvement on t learns by evaluating the target policy policy. ⑤ On-policy learning learns from exper performing policy improvement on t learns by evaluating the results of a lon the behavior policy. Submit You have used 1 of 2 attempts ☑ Eligibility traces down weight the corcomputing average Vs) or Q(s,a). ☐ Eligibility traces assign credit to both states. ☑ Eligibility traces assign credit to both states. ☑ Eligibility traces assign credit to both states. ☑ Vou have used 1 of 2 attempts ☑ Correct (1/1 point) Question 7 I/1 point (graded) Which of the following characterizes Q-Learning uses bootstrapping and Q-Learning uses bootstrapping and Q-Learning is an on-policy algorithm, who sharsA is an on-policy algorithm.	whereas off-policy learns from experience forming policy improvement on the target policy improvement on the target policy and the target policy, whereas off-policy learning behavior policy to perform policy improvement on a behavior policy to perform policy improvement on a behavior policy of the target policy, whereas off-policy learning behavior policy to perform policy improvement on a behavior policy and the target policy whereas off-policy algorithm of the state space. Save Show Answer Save Show Answer C All Rights Reserve All Rights Reserve
policy improvement on a target polic by evaluating a target policy and perpolicy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a lon a target policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from experperforming policy improvement on t learns by evaluating the results of a lon the behavior policy. Submit You have used 1 of 2 attempts ✓ Correct (1/1 point) Question 6 If point (graded) Which two statements describe eligibility to expend to action the policy and the correct policy and the correct policy and the correct policy and the correct policy are policy and the correct policy.	rience by evaluating a target policy and the target policy whereas off-policy learning behavior policy to perform policy improvement on the target policy, whereas off-policy learning behavior policy to perform policy improvement on a behavior policy in perform policy improvement on a behavior policy to perform policy improvement on a behavior policy and the target policy improvement on a behavior policy and the target policy and the target policy improvement on a behavior policy and the target policy an
policy improvement on a target polic by evaluating a target policy and per policy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the results of a lon a target policy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the target policy policy. On-policy learning learns from exper performing policy improvement on t learns by evaluating the results of a lon the behavior policy. Submit You have used 1 of 2 attempts Correct (1/1 point) Question 6 It ligibility traces down weight the cordomputing average Vs) or Q(s,a). Eligibility traces assign credit to action learns by evaluating average Vs) or Q(s,a). Eligibility traces assign credit to both states. Submit You have used 1 of 2 attempts Correct (1/1 point) Question 7 If point (graded) Which of the following characterizes Q-Learning uses bootstrapping and Q-Learning uses bootstrapping and Q-Learning is an on-policy algorithm, whee SARSA uses bootstrapping and Q-Learning is an on-policy algorithm, whee Submit You have used 1 of 2 attempts Correct (1/1 point) Correct (1/1 point) Correct (1/1 point)	ey, whereas off-policy learns from experience forming policy improvement on the target policy improvement on the target policy improvement on the target policy and the target policy whereas off-policy learning behavior policy to perform policy improvement on a behavior policy and the target policy, whereas off-policy learning behavior policy improvement on a behavior policy to perform policy improvement on a behavior policy algorithm on the most frequently visited and last visited to exploration of the state space. Save Show Answer Apply Store Save Show Answer Apply Store Save Show Answer Apply Store Save Show Answer Apply Store Save Show Answer
policy improvement on a target polic by evaluating a target policy and per policy. ② On-policy learning learns from exper performing policy improvement on a learns by evaluating the results of a learns by evaluating the target policy policy. ③ On-policy learning learns from exper performing policy improvement on a learns by evaluating the target policy policy. ⑤ On-policy learning learns from exper performing policy improvement on a learns by evaluating the results of a learns by	sy, whereas off-policy learns from experience forming policy improvement on the target forming policy improvement on the target policy and the target policy whereas off-policy learning behavior policy to perform policy improvement on a behavior policy perform policy improvement on a behavior policy perform policy improvement on a behavior policy behavior policy to perform policy improvement on a behavior policy of the behavior policy algorithm on the most frequently visited and last visited to specification of the state space. Save Show Answer **Application of the state space.** **Application