

## Reference – Shown-as-Correct Answers

July 26, 2025

### Q1

Answer presented as correct: A) Proximal Policy Optimization (PPO)■(

*Reason: Clearly aligns with how the concept is described in the question.*

### Q2

Answer presented as correct: B) Observations lack some hidden variables■(

*Reason: Clearly aligns with how the concept is described in the question.*

### Q3

Answer presented as correct: D) Model-based planning

*Reason: Clearly aligns with how the concept is described in the question.*

### Q4

Answer presented as correct: B) Action-value minus state-value■(

*Reason: Clearly aligns with how the concept is described in the question.*

### Q5

Answer presented as correct: B) Target networks■(

*Reason: Clearly aligns with how the concept is described in the question.*

### Q6

Answer presented as correct: D) Deterministic policies

*Reason: Clearly aligns with how the concept is described in the question.*

### Q7

Answer presented as correct: D) Hyper-parameter tuning■2

*Reason: Clearly aligns with how the concept is described in the question.*