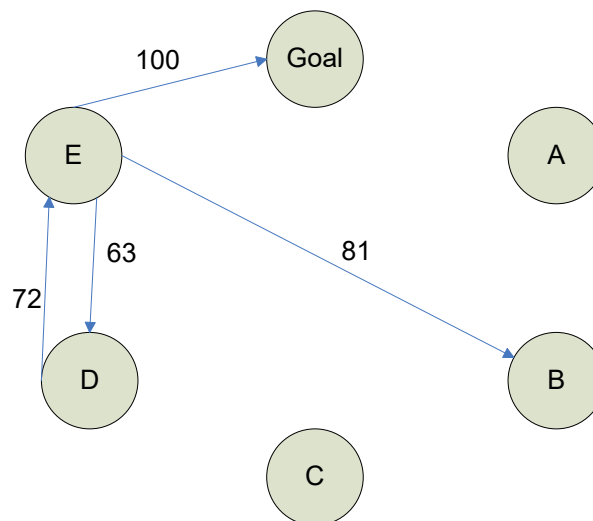


HW Chap. 18

Reinforcement Learning

1. In Reinforcement Learning, what is the difference between the *reward* function and the *value* function?
2. The figure below shows the current state of a system and the relevant Q values that have been hypothesized so far. The learner is currently in state D, and has decided to take the action that moves to state E. Assume the reward upon entering each state is 0, the learning rate η is 0.5, and discount factor γ is 0.9. Show how the state diagram changes based upon Q-Learning.



3. Temporal Difference learning uses something called γ (the “discount factor”), that typically ranges from 0..1. Describe the effect of using the values 0, 0.5, and 1 for γ .