

- (1) Approximate transition probabilities and rewards
- (2) forall  $s \in S$  and  $a \in A$  combinations:
  - (2)  $Q_{s,a} = p'_{s,a}(r + \gamma \max_{a' \in A} Q_{s',a'})$
- (3) Repeat step (1) - (2)