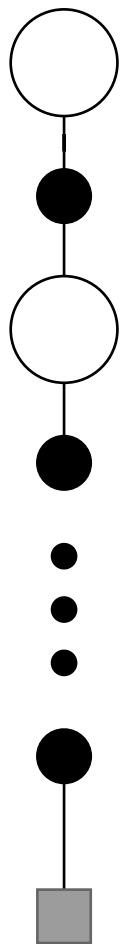
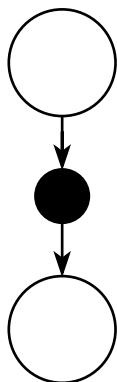


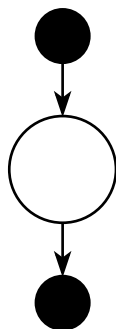
MC



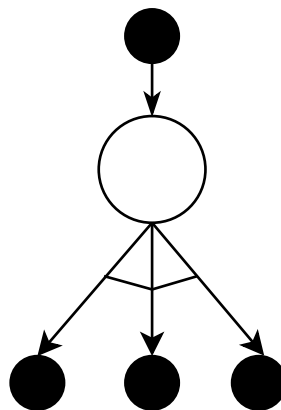
TD(0)



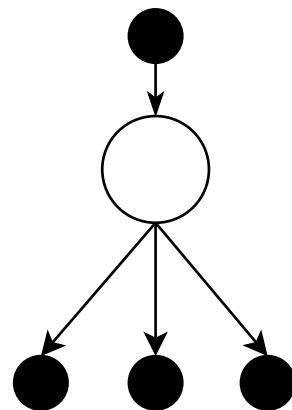
Sarsa



Q-learning



Expected Sarsa



terminal state