EX4

Monte Carlo Methods vs Dynamic Programming

1. MC is way faster for large environments and for DP it is required to know the transition function p(s’,r|s,a), for MC it is not.
2. Blackjack as in exercise two of this sheet.

Programming:

Policy Evalulation with hit at 20: Ein Bild, das Text, Diagramm, Design enthält.

Automatisch generierte Beschreibung  
Trained Policy: It was trained until it does not change anymore. The expected reward is -0.19