

Assignment5: Markov Decision Procedures

Dmitrii, Maksimov
dmitrii.maksimov@fau.de
ko65bexp

Ilia, Dudnik
ilia.dudnik@fau.de
ex69ahum

Aleksandr, Korneev
aleksandr.korneev@fau.de
uw44ylyz

June 19, 2022

Exercise 5.1 (Bellman Equation)

State the Bellman Equation and explain every symbol in the equation and what the equation is used for and how.

$$U(s) = R(s) + \gamma \cdot \max_{a \in A(s)} \sum_{s'} U(s') \cdot P(s'|s, a)$$

, where s - a state, U - expected sum of reward, R - current reward, γ - discount factor, $A(s)$ - set of possible actions in state s , a - action, s' - possible state after the action a .

Bellman Equation is used for updating the utility for a state s . We have to assign some utility at each node at the beginning and then update utility for each state until convergency. After that, we can choose policy as MEU.