Reinforcement Learning - Monsoon 2022
IIITD
Assignment 1
Ojus Singhal | 2020094 | ojus20094@iiitd.ac.in

Answer 4

s	a	s'	r	$p(s', r\|s, a)$
high	search	high	r_{search}	α
high	search	low	r_{search}	$1-\alpha$
low	search	high	-3	$1-\beta$
low	search	low	r_{search}	β
high	wait	high	r_{wait}	1
low	wait	low	r_{wait}	1
low	recharge	high	0	1

Here I assumed that the rewards are deterministic rather than stochastic, since no distribution in provided for the rewards. So for all possible state transitions with non-zero probabilities, we just give that reward with probability 1, and values for $p(s', r|s, a)$ become the probabilities of $p(s'|s, a)$ for the corresponding state transition.