Name – Yunika Upadhayaya
ID – 1001631183
Assignment 7 – Written

Task – 1 (Output):

a) value_iteration('environment2.txt', -0.04, 1, 20)

utilities:
0.812 0.868 0.918 1.000
0.762 0.000 0.660 -1.000
0.705 0.655 0.611 0.387

policy:
> > > *o*
^ *x* ^ *o*
^ < < <

b) value_iteration('environment2.txt', -0.04, 0.9, 20)

utilities:
0.509 0.650 0.795 1.000
0.399 0.000 0.486 -1.000
0.296 0.254 0.345 0.130

policy:
> > > *o*
^ *x* ^ *o*
^ > ^ <

Task – 2:

Task – 3:

Part a:

Considering optimal action as ‘up’:
$U_h((2, 2), (3, 2)) = - 0.04 + 0.9 * 1 = 0.86$
$U_h((2, 2), (2, 3)) = - 0.04 + 0.9 * U(2, 2)$
$U_h((2, 2), (2, 1)) = - 0.04 + 0.9 * U(2, 2)$

Now,
U (2, 2) = 0.8 * 0.86 + 0.1 * (- 0.04 + 0.9 * U (2, 2)) + 0.1 * (- 0.04 + 0.9 * U (2, 2))
U (2, 2) = 0.68 + 0.18 U (2, 2)
U (2, 2) = 0.68 / 0.8 = 0.829

Considering optimal action as 'down':
U_h ((2, 2), (1, 2)) = - 0.04 + 0.9 * -1 = - 0.94
U_h ((2, 2), (2, 3)) = - 0.04 + 0.9 * U (2, 2)
U_h ((2, 2), (2, 1)) = - 0.04 + 0.9 * U (2, 2)

Now,
U (2, 2) = - 0.94 * 0.8 + 0.2 * (-0.04 + 0.9 U (2,2))
U (2, 2) = - 0.76 / 0.82 = - 0.93

Considering optimal action as 'left':
U_h ((2, 2), (3, 2)) = - 0.04 + 0.9 * 1 = 0.86
U_h ((2, 2), (2, 1)) = - 0.04 + 0.9 * U (2, 2)
U_h ((2, 2), (1, 2)) = - 0.04 + 0.9 * -1 = - 0.94

Now,
U (2, 2) = 0.8 * (-0.04 + 0.9 U (2, 2)) + 0.1 * - 0.94 + 0.1 * 0.86
U (2, 2) = - 0.04/0.28 = -0.143

Considering optimal action as 'right':
U_h ((2, 2), (3, 1)) = - 0.04 + 0.9 * 1 = 0.86
U_h ((2, 2), (2, 3)) = - 0.04 + 0.9 * U (2, 2)
U_h ((2, 2), (1, 2)) = - 0.04 + 0.9 * - 1 = -0.94

Now,
U (2, 2) = 0.8 * (- 0.04 + 0.9 U (2, 2)) + 0.1 * - 0.94 + 0.1 *0.86
U (2, 2) = - 0.04 / 0.28 = - 0.143

Thus, the maximum value obtained is from optimal action as 'up': **U (2, 2) = 0.829**