Contributions:

Jonathan Avila: I worked in the problem of Monte Carlo trees, I implemented all the code for such problem and evaluated the code for Oscar’s work. I worked with Oscar to develop the Q-learning code.

Oscar Galindo: I worked in the problem of Value Iteration, I implemented all the code for such problem and evaluated the code for Jonathan’s work. I worked with Jonathan to develop the Q-learning code.

Outputs per program:

Monte Carlo output:

Ep 1: [RU 8p ] → S → -1 → [RD 10p] → P → +2 → [RD 8a ] → R → +0 → [RD 10a] → P → +4 → [11am class begins] r=5

Ep 2: [RU 8p ] → P → +2 → [TU 10p] → P → +2 → [RU 10a] → S → +0 → [11am class begins] r=4

Ep 3: [RU 8p ] → S → -1 → [RD 10p] → P → +2 → [RD 8a ] → P → +2 → [TD 10a] → S → +3 → [11am class begins] r=6

Ep 4: [RU 8p ] → S → -1 → [RD 10p] → P → +2 → [RD 10a] → P → +4 → [11am class begins] r=5

Ep 5: [RU 8p ] → P → +2 → [TU 10p] → R → +0 → [RU 8a ] → P → +2 → [TU 10a] → P → -1 → [11am class begins] r=3

Ep 6: [RU 8p ] → P → +2 → [TU 10p] → R → +0 → [RU 8a ] → R → +0 → [RU 10a] → R → +0 → [11am class begins] r=2

Ep 7: [RU 8p ] → R → +0 → [RU 10p] → R → +0 → [RU 8a ] → S → -1 → [RD 10a] → R → +4 → [11am class begins] r=3

Ep 8: [RU 8p ] → S → -1 → [RD 10p] → P → +2 → [RD 10a] → R → +4 → [11am class begins] r=5

Ep 9: [RU 8p ] → P → +2 → [TU 10p] → R → +0 → [RU 8a ] → P → +2 → [TU 10a] → R → -1 → [11am class begins] r=3

Ep 10: [RU 8p ] → P → +2 → [TU 10p] → R → +0 → [RU 8a ] → R → +0 → [RU 10a] → P → +0 → [11am class begins] r=2

Ep 11: [RU 8p ] → P → +2 → [TU 10p] → P → +2 → [RU 10a] → S → +0 → [11am class begins] r=4

Ep 12: [RU 8p ] → R → +0 → [RU 10p] → R → +0 → [RU 8a ] → P → +2 → [TU 10a] → R → -1 → [11am class begins] r=1

Ep 13: [RU 8p ] → S → -1 → [RD 10p] → R → +0 → [RD 8a ] → P → +2 → [TD 10a] → P → +3 → [11am class begins] r=4

Ep 14: [RU 8p ] → S → -1 → [RD 10p] → R → +0 → [RD 8a ] → P → +2 → [TD 10a] → P → +3 → [11am class begins] r=4

Ep 15: [RU 8p ] → P → +2 → [TU 10p] → P → +2 → [RU 10a] → S → +0 → [11am class begins] r=4

Ep 16: [RU 8p ] → R → +0 → [RU 10p] → S → -1 → [RD 8a ] → P → +2 → [TD 10a] → R → +3 → [11am class begins] r=4

Ep 17: [RU 8p ] → R → +0 → [RU 10p] → P → +2 → [RU 8a ] → P → +2 → [TU 10a] → R → -1 → [11am class begins] r=3

Ep 18: [RU 8p ] → P → +2 → [TU 10p] → R → +0 → [RU 8a ] → S → -1 → [RD 10a] → R → +4 → [11am class begins] r=5

Ep 19: [RU 8p ] → P → +2 → [TU 10p] → P → +2 → [RU 10a] → P → +0 → [11am class begins] r=4

Ep 20: [RU 8p ] → P → +2 → [TU 10p] → P → +2 → [RU 10a] → P → +0 → [11am class begins] r=4

Ep 21: [RU 8p ] → P → +2 → [TU 10p] → P → +2 → [RU 10a] → P → +0 → [11am class begins] r=4

Ep 22: [RU 8p ] → R → +0 → [RU 10p] → P → +2 → [RU 8a ] → R → +0 → [RU 10a] → P → +0 → [11am class begins] r=2

Ep 23: [RU 8p ] → R → +0 → [RU 10p] → P → +2 → [RU 8a ] → P → +2 → [TU 10a] → P → -1 → [11am class begins] r=3

Ep 24: [RU 8p ] → R → +0 → [RU 10p] → P → +2 → [RU 8a ] → P → +2 → [TU 10a] → R → -1 → [11am class begins] r=3

Ep 25: [RU 8p ] → S → -1 → [RD 10p] → P → +2 → [RD 8a ] → R → +0 → [RD 10a] → P → +4 → [11am class begins] r=5

Ep 26: [RU 8p ] → S → -1 → [RD 10p] → R → +0 → [RD 8a ] → R → +0 → [RD 10a] → R → +4 → [11am class begins] r=3

Ep 27: [RU 8p ] → R → +0 → [RU 10p] → P → +2 → [RU 8a ] → P → +2 → [TU 10a] → S → -1 → [11am class begins] r=3

Ep 28: [RU 8p ] → S → -1 → [RD 10p] → P → +2 → [RD 8a ] → P → +2 → [TD 10a] → S → +3 → [11am class begins] r=6

Ep 29: [RU 8p ] → R → +0 → [RU 10p] → S → -1 → [RD 8a ] → P → +2 → [TD 10a] → P → +3 → [11am class begins] r=4

Ep 30: [RU 8p ] → R → +0 → [RU 10p] → S → -1 → [RD 8a ] → R → +0 → [RD 10a] → P → +4 → [11am class begins] r=3

Ep 31: [RU 8p ] → R → +0 → [RU 10p] → P → +2 → [RU 8a ] → P → +2 → [TU 10a] → R → -1 → [11am class begins] r=3

Ep 32: [RU 8p ] → R → +0 → [RU 10p] → R → +0 → [RU 8a ] → P → +2 → [TU 10a] → P → -1 → [11am class begins] r=1

Ep 33: [RU 8p ] → S → -1 → [RD 10p] → R → +0 → [RD 8a ] → R → +0 → [RD 10a] → R → +4 → [11am class begins] r=3

Ep 34: [RU 8p ] → R → +0 → [RU 10p] → P → +2 → [RU 8a ] → R → +0 → [RU 10a] → P → +0 → [11am class begins] r=2

Ep 35: [RU 8p ] → P → +2 → [TU 10p] → P → +2 → [RU 10a] → S → +0 → [11am class begins] r=4

Ep 36: [RU 8p ] → R → +0 → [RU 10p] → R → +0 → [RU 8a ] → P → +2 → [TU 10a] → S → -1 → [11am class begins] r=1

Ep 37: [RU 8p ] → R → +0 → [RU 10p] → R → +0 → [RU 8a ] → S → -1 → [RD 10a] → P → +4 → [11am class begins] r=3

Ep 38: [RU 8p ] → P → +2 → [TU 10p] → P → +2 → [RU 10a] → R → +0 → [11am class begins] r=4

Ep 39: [RU 8p ] → R → +0 → [RU 10p] → P → +2 → [RU 10a] → S → +0 → [11am class begins] r=2

Ep 40: [RU 8p ] → S → -1 → [RD 10p] → P → +2 → [RD 10a] → S → +4 → [11am class begins] r=5

Ep 41: [RU 8p ] → P → +2 → [TU 10p] → P → +2 → [RU 10a] → R → +0 → [11am class begins] r=4

Ep 42: [RU 8p ] → R → +0 → [RU 10p] → P → +2 → [RU 8a ] → R → +0 → [RU 10a] → R → +0 → [11am class begins] r=2

Ep 43: [RU 8p ] → P → +2 → [TU 10p] → R → +0 → [RU 8a ] → P → +2 → [TU 10a] → P → -1 → [11am class begins] r=3

Ep 44: [RU 8p ] → P → +2 → [TU 10p] → P → +2 → [RU 10a] → P → +0 → [11am class begins] r=4

Ep 45: [RU 8p ] → P → +2 → [TU 10p] → R → +0 → [RU 8a ] → P → +2 → [TU 10a] → P → -1 → [11am class begins] r=3

Ep 46: [RU 8p ] → S → -1 → [RD 10p] → P → +2 → [RD 10a] → S → +4 → [11am class begins] r=5

Ep 47: [RU 8p ] → R → +0 → [RU 10p] → S → -1 → [RD 8a ] → P → +2 → [TD 10a] → S → +3 → [11am class begins] r=4

Ep 48: [RU 8p ] → R → +0 → [RU 10p] → P → +2 → [RU 10a] → S → +0 → [11am class begins] r=2

Ep 49: [RU 8p ] → S → -1 → [RD 10p] → R → +0 → [RD 8a ] → R → +0 → [RD 10a] → P → +4 → [11am class begins] r=3

Ep 50: [RU 8p ] → R → +0 → [RU 10p] → S → -1 → [RD 8a ] → P → +2 → [TD 10a] → P → +3 → [11am class begins] r=4

StateVals:

[RU 8p ] +3.332

[TU 10p] +3.042

[RU 10p] +2.345

[RD 10p] +3.271

[RU 8a ] +2.201

[RD 8a ] +3.040

[TU 10a] +1.767

[RU 10a] +2.571

[RD 10a] +2.958

[TD 10a] +2.520

[11am class begins] +3.332

AvgReward:+3.460

Value Iteration Output:

--------------------------------------------

Finalized iterations. Ran for 5 iteration, optimal policy by state ['S', 'R', 'S', 'P', 'S', 'P', 'R', 'R', 'R', 'R']

Values: [5.3757485, 2.9304, 3.9203, 6.44015, 2.96, 4.97, -1.0, 0.0, 4.0, 3.0, 0.0]

Q-learning Output:

--------------------------------------------

Number of episodes is: 409

Q values: [[6.9410973718555615, 3.3433191271201226, 6.473342237124173], [2.768790284919669, 4.518087997506665], [2.155832114216908, 3.4628092983944576, 2.633236597249897], [5.146228674389209, 7.224408138741257], [1.1971521012867223, 1.9383268258948054, 2.9056432159365886], [4.941587440549737, 3.280411052018597], [-0.3848805222779272, -0.18769676278925793, -0.24771643368071342], [1.3231968934241656, 0.5845692994451522, 1.5961205980351851], [5.0482052136759, 4.838190876522985, 4.935880158440396], [1.7835608222335013, 1.6936275648259471, 2.0611110149377536], [0]]