In the name of God

Assignment 5 Solution

Neural Networks: Fall 1400, Dr. Mozayani

Ghazaleh Mahmoudi

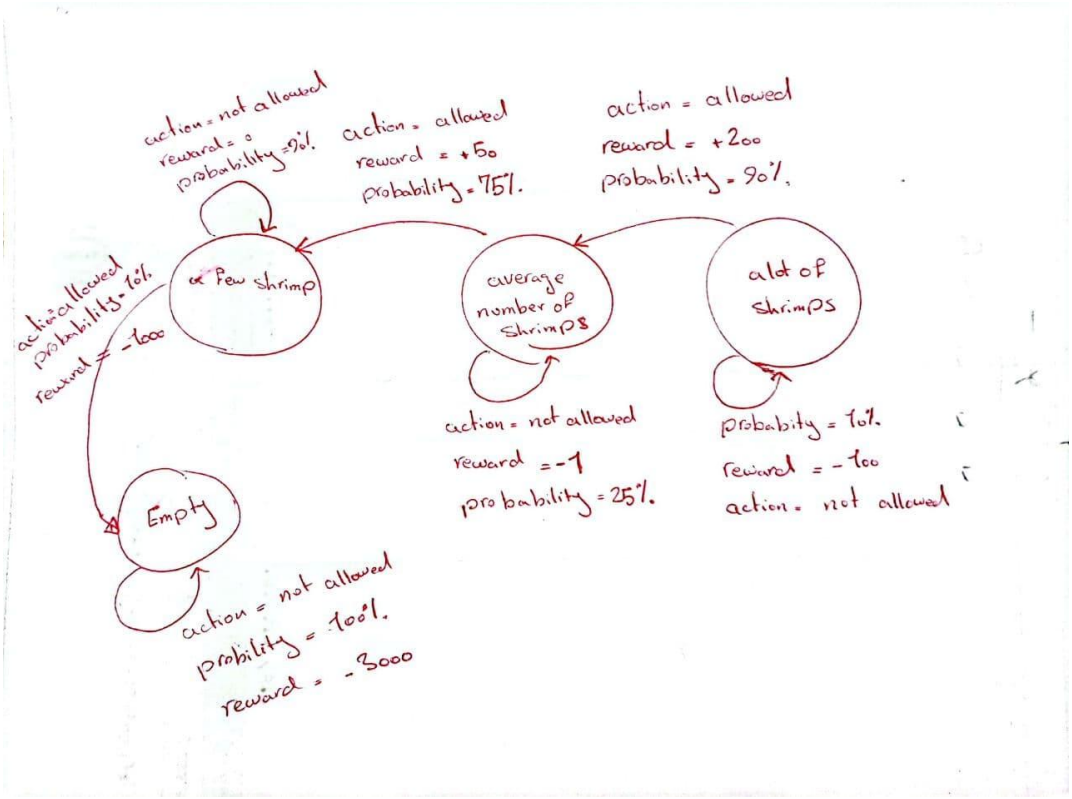
400722156

23 January 2022

Problem 1

- (a)
  - States.
    - No shrimp/Empty.
    - A few shrimp.
    - Average number of shrimp.
    - Lots of shrimp.
  - Actions.
    - sale of shrimps is **allowed**.
    - sale of shrimps is **not allowed**.

- State transition probability + Reward



- Reward
  - **The lower the number of shrimp, the lower the reward.**
  - **The policy is arranged so that the number of shrimp is not too small.**

- (b)
  - **States.**
    - Falling in the pit (terminal state).
    - dying by enemies (terminal state).
    - piranha plate on green pipe (terminal state).
    - end flag (terminal state).
    - **continue playing (Mario is still alive).**
  - **Actions**
    - Jump, move forward, fast move forward, slow move forward.
  - **State transition probability.**
    - Possibility (cur_state, next_state, [action])
    - Possibility (Mario is still alive , Falling in the pit) : 5%
      - Possibility (Mario is still alive , Falling in the pit, fast move forward) = 2.4%
      - Possibility (Mario is still alive , Falling in the pit, slow move forward) = 2.4%

- Possibility (Mario is still alive , Falling in the pit, other actions): 0.2%

■ Possibility(Mario is still alive , dying by enemies) : 10%

■ Possibility (Mario is still alive, piranha plate on green pipe): 4%

- Possibility (Mario is still alive, piranha plate on green pipe, fast move forward): 3.9%
- Possibility (Mario is still alive, piranha plate on green pipe, other actions): 0.1%

■ Possibility (Mario is still alive , end flag) : 1%

■ Possibility (Mario is still alive, Mario is still alive) : 80%

■ **Other state transition Possibility is 0;**

○ **Reward**

■ Falling in the pit : -90

■ dying by enemies : -90

■ piranha plate on green pipe : -90

■ end flag : +150

■ **continue playing (Mario is still alive)** : -1

Problem 2

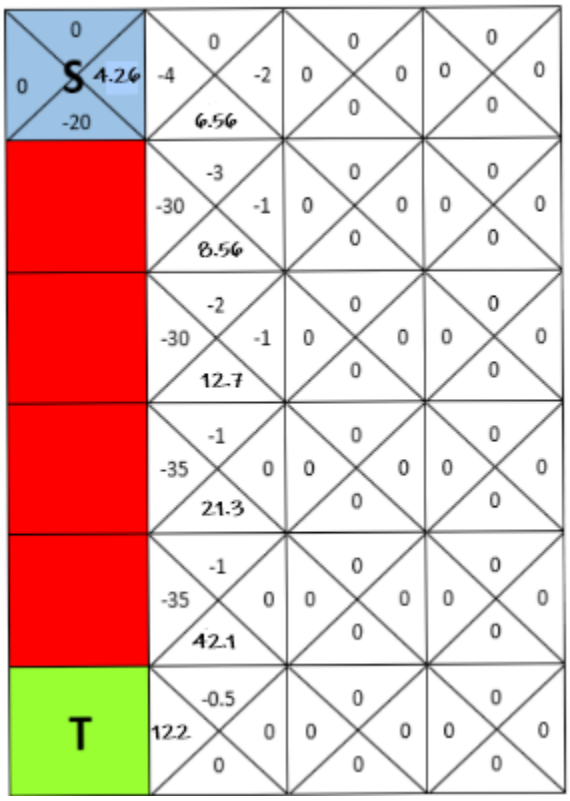
$$Q(a,i) \leftarrow Q(a,i) + \alpha(R(i) + \gamma \max_{a'} Q(a',j) - Q(a,i))$$

episode 1: Right, Down, Down, Down, Down, Down, Left.

- Q(S, Right) = 3 + 0.9 × (-1 + 0.8 × (6-3)) = 4.26
  - +3 -> 4.26
- Q(S, Down) = 6 + 0.9 × (-1 + 0.8 × (8-6)) = 6.56
  - +6 -> 6.56
- Q(S, Down) = 8 + 0.9 × (-1 + 0.8 × (10-8)) = 8.54
  - +8 -> 8.54
- Q(S, Down) = 10 + 0.9 × (-1 + 0.8 × (15-10)) = 12.7
  - +10 -> 12.7
- Q(S, Down) = 15 + 0.9 × (-1 + 0.8 × (25-15)) = 21.3
  - +15 -> 21.3
- Q(S, Down) = 25 + 0.9 × (-1 + 0.8 × (50-25)) = 42.1

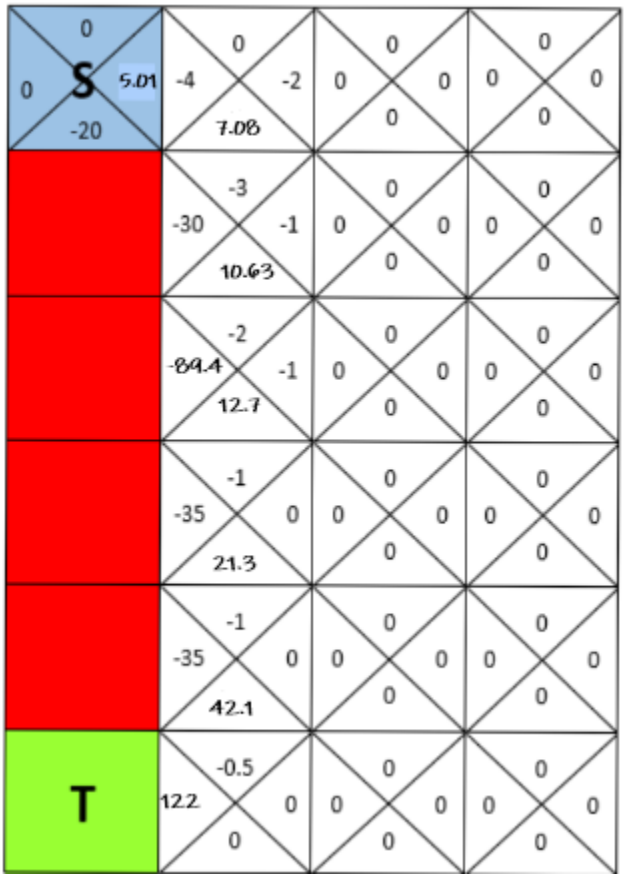
- +25 -> 42.1

- Q(S, Left) = 50 + 0.9 × (120 + 0.8 × (0 - 50)) = 122
  - +50 -> 122



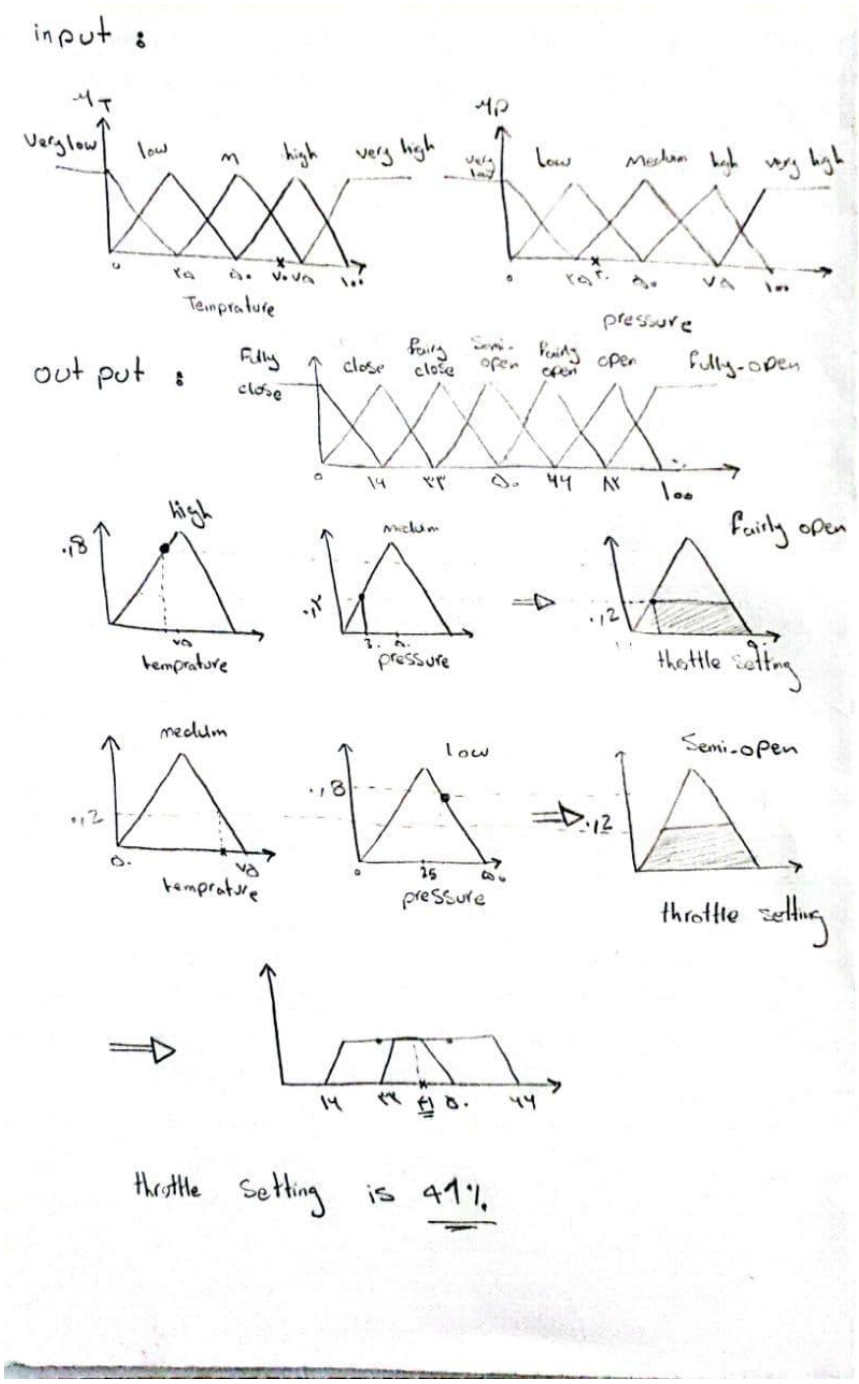
episode 2: Right, Down, Down, Left.

- Q(S, Right) = 4.26 + 0.9 × (-1 + 0.8 × (6.56 - 4.26)) = 5.01
  - 4.26 -> 5.01
- Q(S, Down) = 6.56 + 0.9 × (-1 + 0.8 × (8.54 - 6.56)) = 7.08
  - 6.54 -> 7.08
- Q(S, Down) = 8.54 + 0.9 × (-1 + 0.8 × (12.7 - 8.54)) = 10.63
  - 8.54 -> 10.63
- Q(S, Left) = -30 + 0.9 × (-90 + 0.8 × (-1 - (-30))) = -89.4
  - -30 -> -89.4



Problem 4

- Rules (5*5*7 are available):
  - If temperature is high and pressure is medium then the tottle is fairly open.
  - If temperature is medium and pressure is low then the tottle is semi-open.
  -



throttle setting is 41%.

Problem 5 (Bonus)

- Please refer to the GhazalehMmahmoudi_HW5.ipynb for the complete code