

- Q1.** B (['u','d','l','r'])
- Q2.** C (25)
- Q3.** D (Not Sure)
- Q4.** D (4)
- Q5.1** B (state_action = self.q_table.loc[observation, :]
 action = np.random.choice(state_action[state_action ==
 np.max(state_action)].index))
- Q5.2** C (state_action = self.q_table.loc[observation, :]
 action = np.random.choice(state_action[state_action ==
 np.min(state_action)].index))
- Q6.** A (Always the same, no matter how many training episodes there are)
- Q7.** F (oscillates all the time without any increasing/decreasing trend)
- Q8.** D (Not Sure (who knows?))
- Q9.** See Below

