# RL Note

lijialu1004

August 2023

## 1 Terminologies

| Basic Concepts | | |
|---|---|---|
| Term | Math | Explanation |
| State | $s_i$ | |
| State space | $S$ | |
| Action | $a_i$ | |
| Action space of a state | $A(s_i)$ | action |
| Policy | $\pi(a_j\|s_i)$ | the action $a_j$ taken in state $s$ |
| reward | $r$ | positive=encouragement; negative=pu |
| return | | sum of all reward |
| Discount rate | $\gamma$ | |
| Discounted return | $G_t$ | $G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} +$ |
| Episode /Trail | | agent stop at terminal state, the resulting trajectory is |
| Bellman Equation | | |
| State value | $v\pi(s) = E[Gt\|St = s]$ | start from state $s$, based on policy $\pi$, represents the "value |
| Bellman Equation | $v = r + \gamma Pv$ | |
| Bellman Optimality Equation | | |
| Value Iteration and Policy Iteration | | |