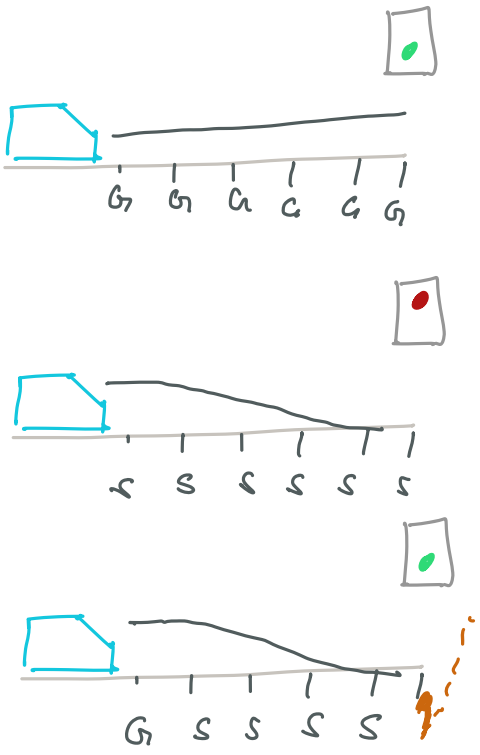
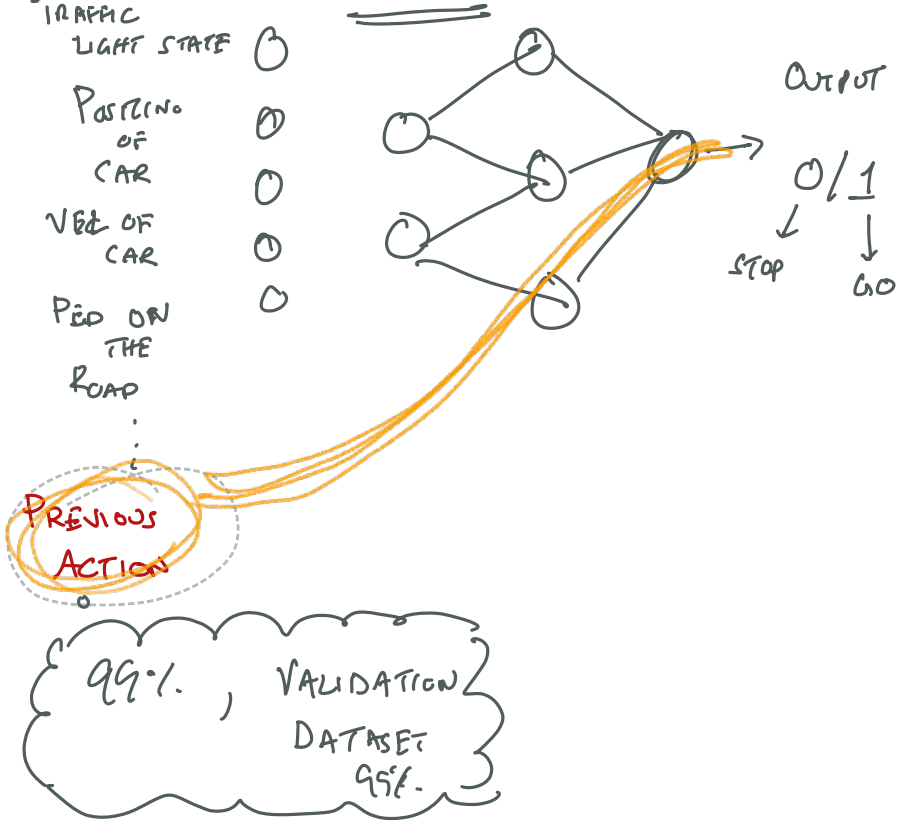
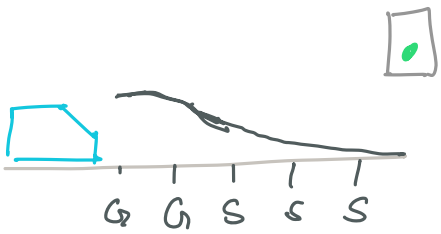
HUMAN DEMONSTRATOR



MODEL



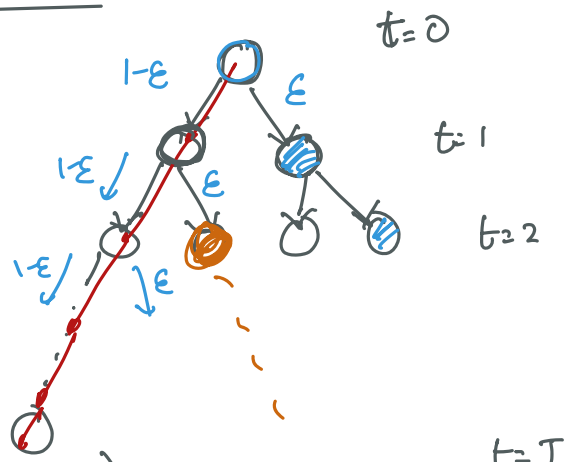
99%, VALIDATION DATASET 95%.



Proof of $O(\varepsilon T^2)$ from BC

$$\mathbb{E}_{\substack{s_0 \sim d_0^t \\ a_t \sim \pi(s_t)}} \sum_{t=1}^{T} c(s_t, a_t)$$

$$\mathbb{E}_{\substack{s_t, a_t \\ \sim d_\pi^t}} \sum c(s_t, a_t) = 0$$



$$J(\pi) - J(\pi^*)$$

$$= \varepsilon \Big(\underbrace{1 + 1 + 1 + 1 + \cdots 1}_{T} \Big)$$

$$+ (1-\varepsilon)\Big(0 + \varepsilon \big(\underbrace{1 + 1 + \cdots 1}_{T-1} \big) \Big)$$

$$+ (1-\varepsilon)(1-\varepsilon)\varepsilon(T-2) + \cdots$$

$$= \varepsilon T + \underbrace{(1-\varepsilon)}_{\leq 1}\varepsilon(T-1) + \underbrace{(1-\varepsilon)^2}_{\leq 1}\varepsilon(T-2) + \cdots$$

$$\leq \varepsilon \left(T + T-1 + T-2 + \cdots \right)$$

$$\leq \varepsilon \frac{T(T+1)}{2}$$

$$O(\varepsilon T^2)$$