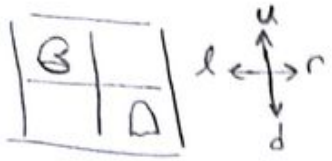
سارا آذرنوش ۹۸۱۷۰۴۱۸

تمرین ۷-۱



۹۰٪
۱۰٪

زنده ماندن +۱

MDP → 3 State

State , Action , transition Probability

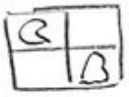
$M = (S, A, R, P, \gamma), 0 < \gamma < 1$

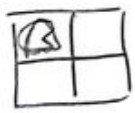
$\forall s \in S : V^{\pi_1}(s) \geq V^{\pi_2}(s) \longrightarrow \pi_1 > \pi_2$, V_π = value Function

V_π

(ا) حداقل یک سیاست وجود دارد که از بقیه سیاست ها بهتر باشد

(ب) deterministic





ghost = g

Paeman = P

$$V_{k+1}^{\pi}(s) \leftarrow \sum_{s'} T(s, \pi(s), s')\left[R(s, \pi(s), s') + \gamma V_k^{\pi}(s')\right]$$