Quiz 90 min Breakout Rooms

Last Time

Online Method

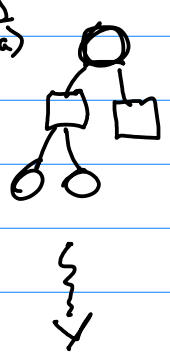
This Time

HW3 - POMDPs.jl

MCTS tips

Continuous state, action

$Q(s,a) + c\sqrt{\frac{\log N(s)}{N(s,a)}}$



For it to work

- Choose c start $2(R_{max} - R_{min})$
- Domain Knowledge
- Rollout Policy / Value Estimate
- Q_0 / N_0

Continuous state, action

$$B[V](s) = \max_a \left(R(s,a) + \gamma \int_S T(s'|s,a) V(s') ds' \right)$$

Linear Dynamics Quadratic Reward

$$s' = T_s s + T_a a + w$$

w zero mean, finite var

$$R(s,a) = s^T R_s s + a^T R_a a$$

Finite Horizon

$$V_h(s) = s^T P_h s + a_h$$

$$\pi^*(s) = -\left(R_a + T_a^T P_h T_a\right)^{-1} T_a^T P_h T_s s = -K_h s$$

discrete time Riccati Eq:

$$P_{h+1} = R_s + T_s^T P_h T_s - \left(T_a^T P_h T_s\right)^T \left(R_a + T_a^T P_h T_a\right)^{-1} \left(T_a^T P_h T_s\right)$$

if $w_i \sim N(0, \Sigma)$ $\quad q_{h+1} = \sum_{i=1}^{h} T_r(\Sigma P_i)$

LQR.

If not Linear Quadratic

- Offline: Approximate Dynamic Prog (ADP)
- Online: - MPC (Model-Predictive)
  - Sparse Tree Search