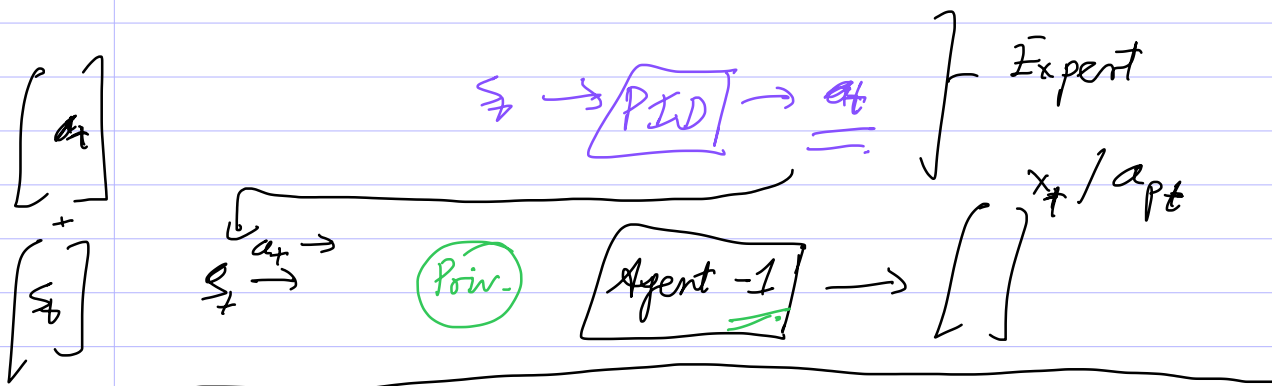
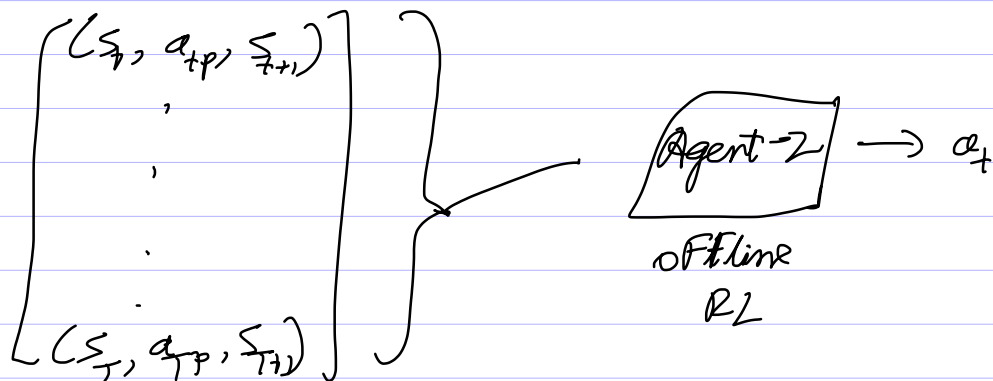


Scheme,

Privileged agent- PID controller



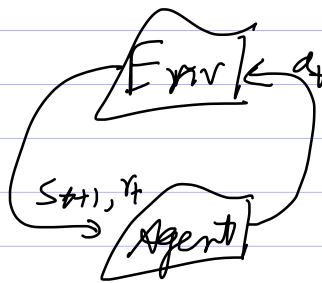
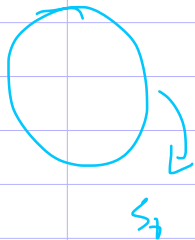
Behavioral
policy is
privileged!!!



Expert

- 1) PID
- 2) LQR
- 3) ILQR
- 4) Model based RL

$$x_t + \sqrt{\alpha_t} x_{t+1}$$



Agent

- 1) QQL
- 2) BCQ
- 3) EMaQ

OFFLINE RL

Learn!

3

xml, C++



1) ResNet 18, ResNet 34

2) 4-6 hours (Appendix A, B)