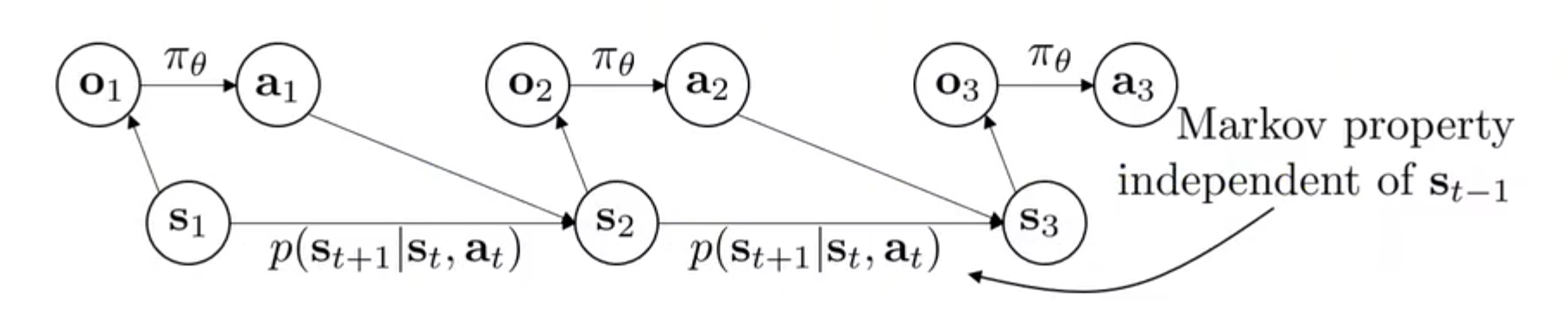
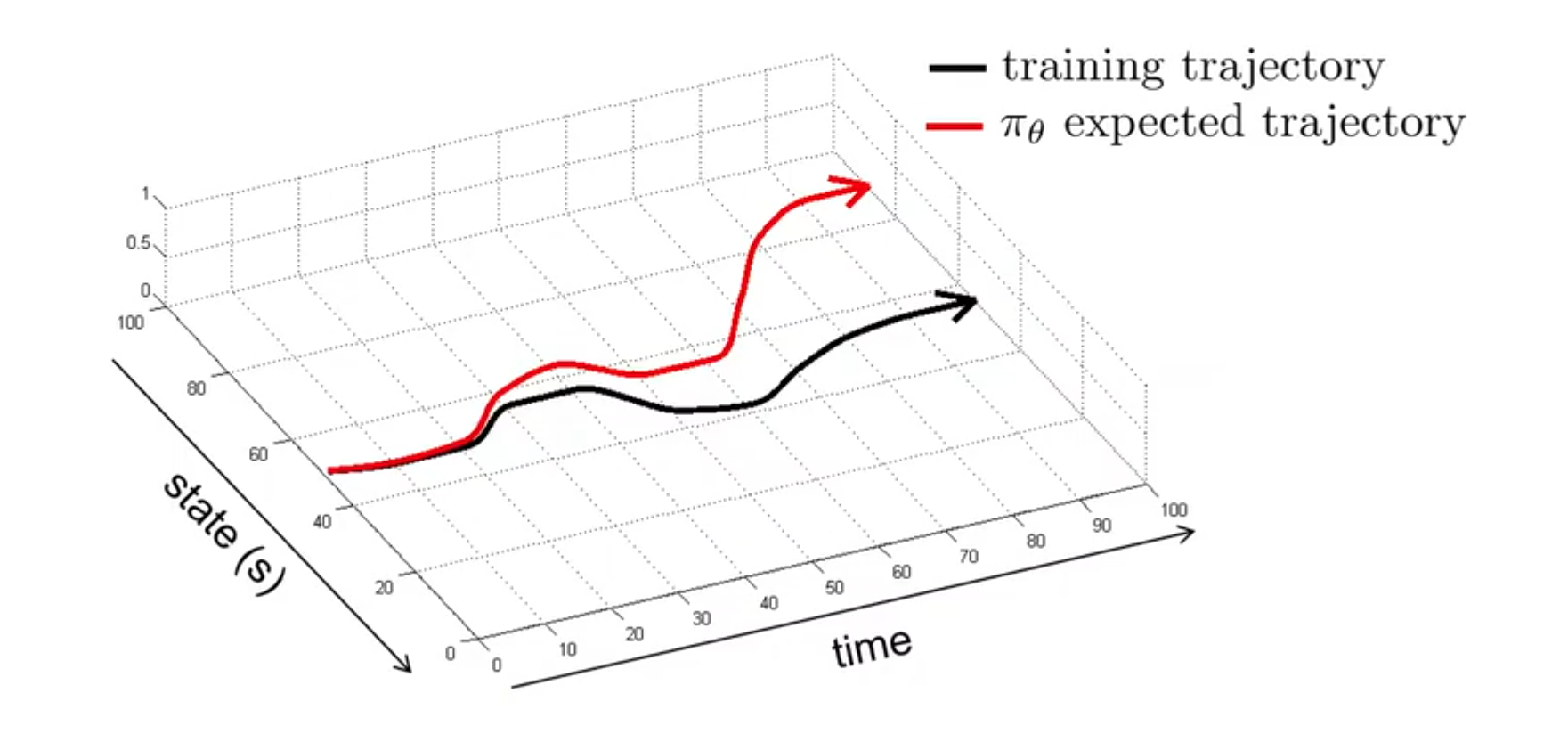
* **Lecture 2 Imitation Learning (Behavior Cloning)**

1. State ≠ Observation ( when the state cannot be fully observed )
2. MDP (Markov Decision Process)



1. The core difference between conventional supervised learning tasks and behavior cloning is independent and identically distributed (IID) assumption. The below is a common problem which can reflect the reason why behavior cloning does not work.



1. The solutions to this problem mainly conclude: collect or augment data, improve the net, multi-task learning, change the algorithm (Dagger).
2. When the task cannot be considered MDP, so the expert naturally cannot be cloned. Maybe you can use temporal information, but it also may lead to casual confusion problem.
3. The limitation of behavior cloning is that it requires a lot of data to train which are probably not easy to get and cannot be better than human.