Imitation learning

Learning to act

- How do we train our policy?

- Imitation learning
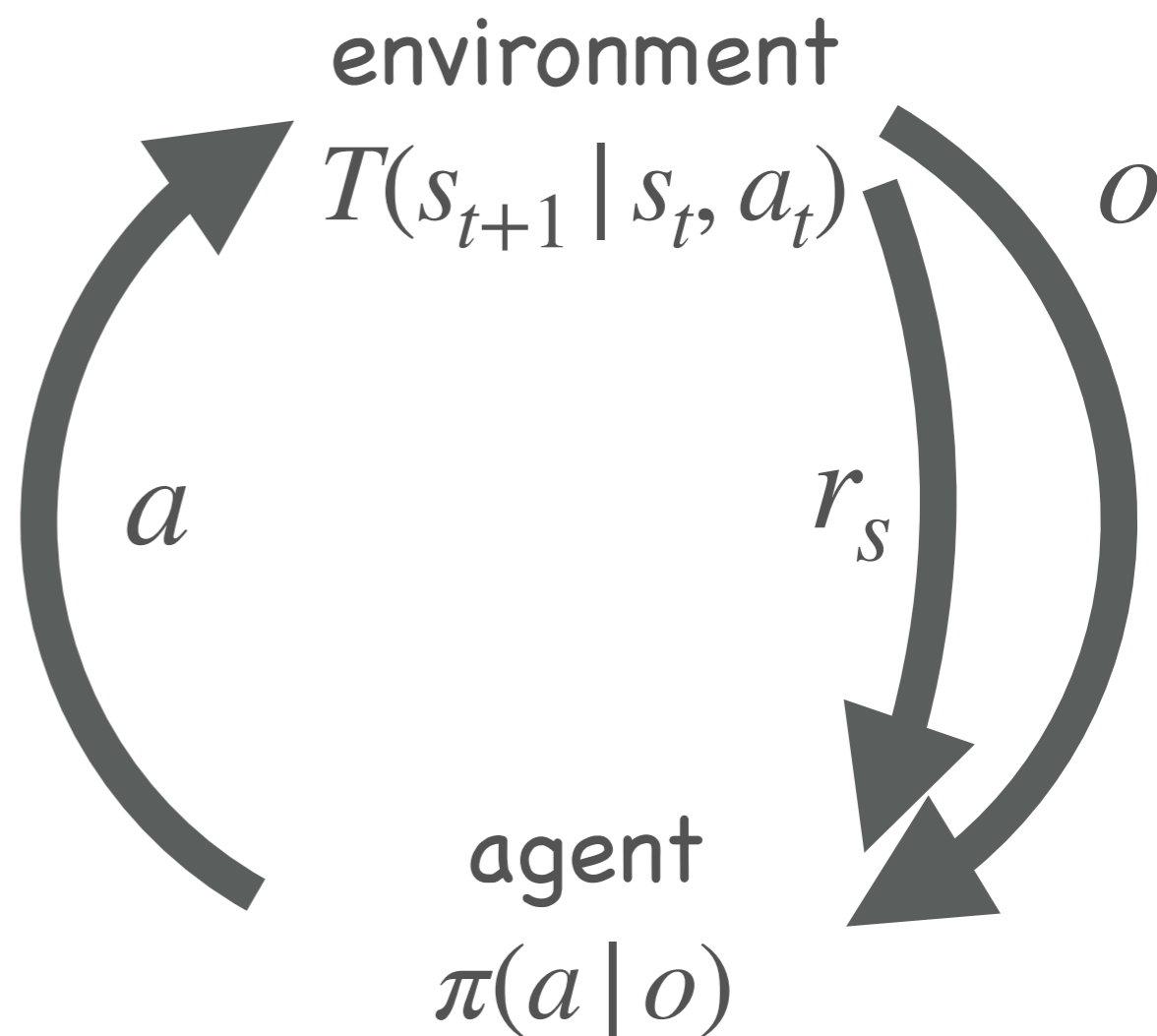
  - Ask an expert

| t=0 | t=1 | t=2 |
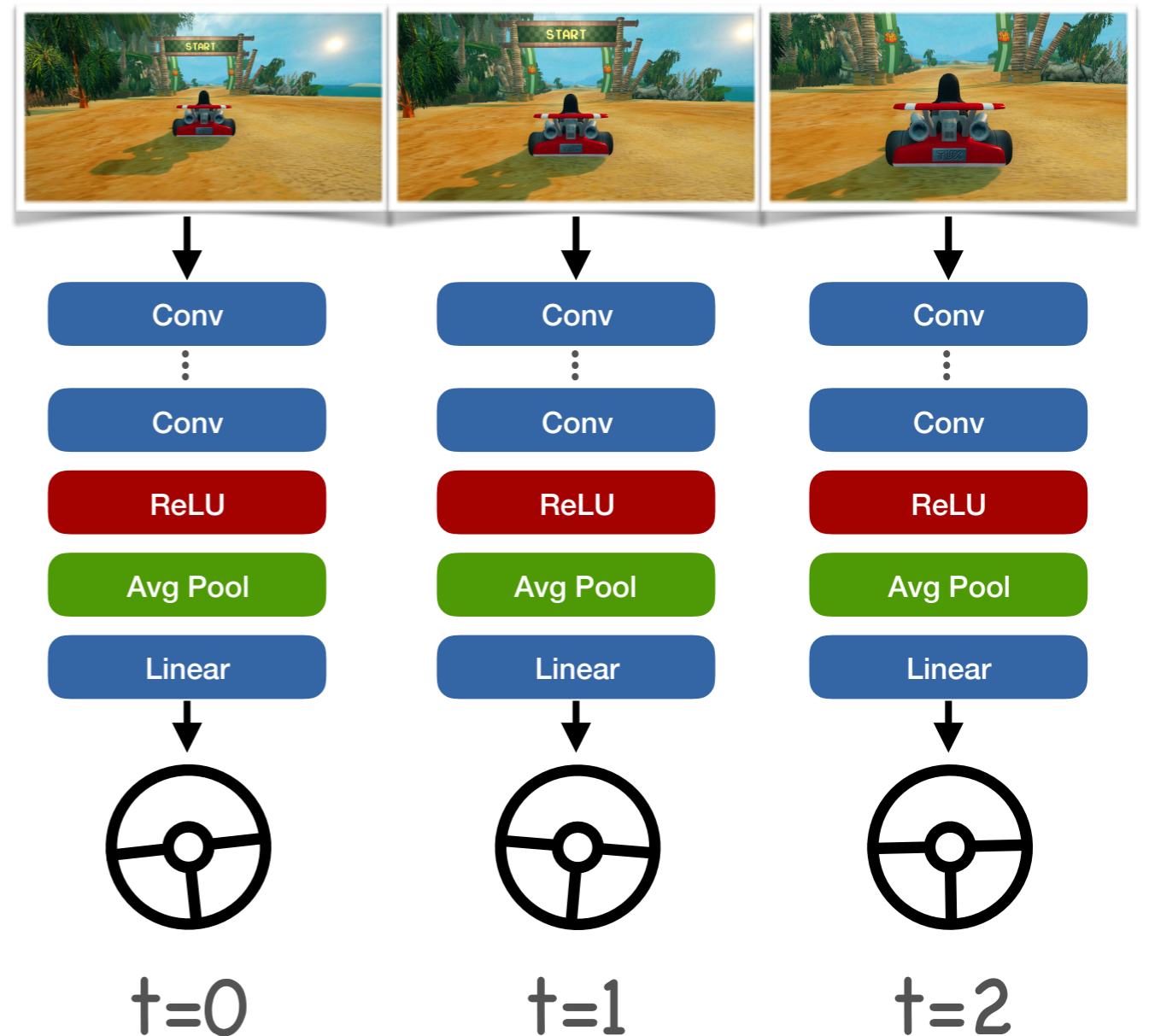
Imitation learning - definition

- Oracle / expert

  - Provides trajectories τ with high return
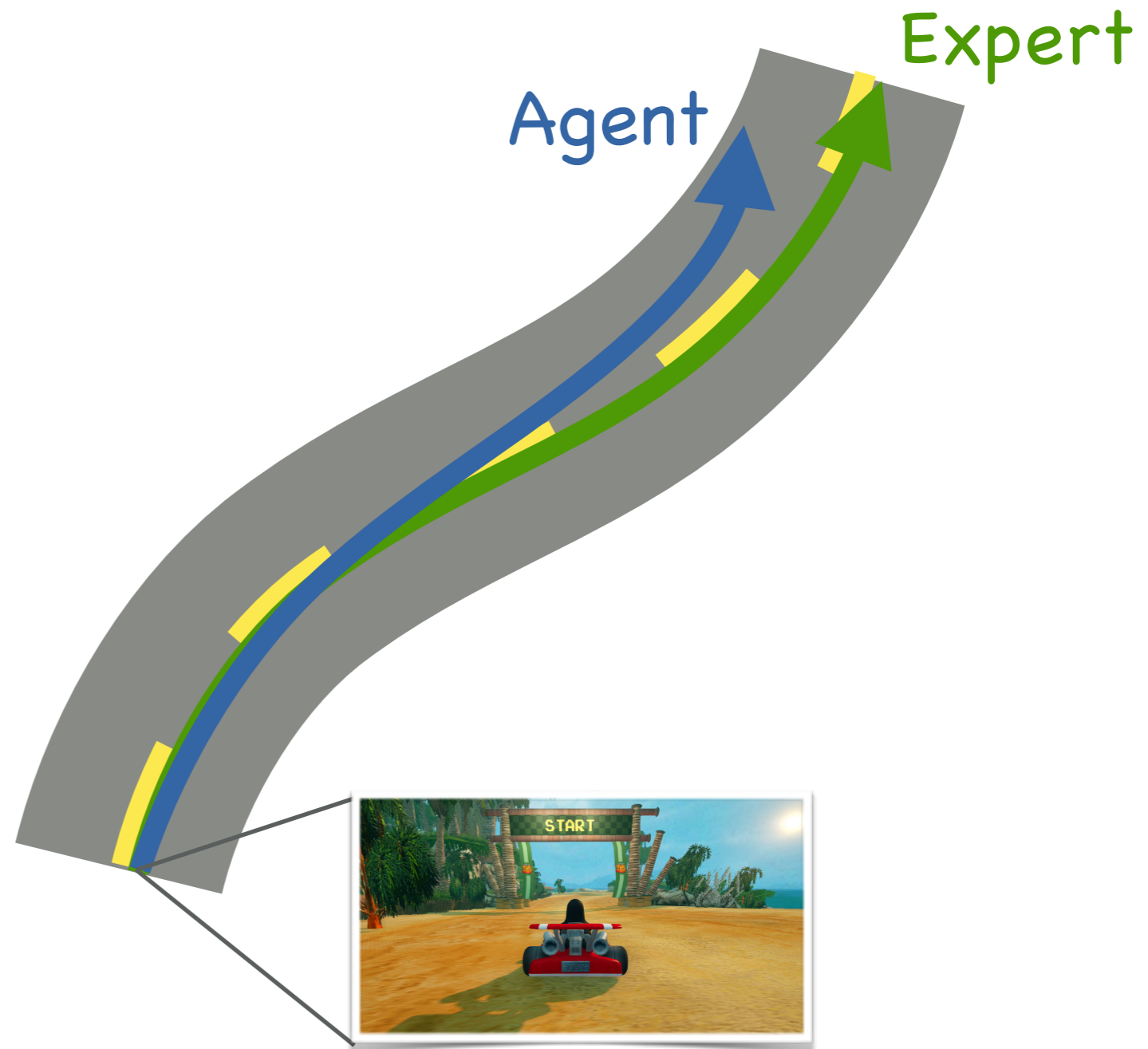
- Policy

  - Supervised learning

  $$\text{maximize}_\pi \, \mathbb{E}_\tau \left[\sum_t \log \pi(a_t, s_t) \right]$$

$$\text{maximize}_\pi \, \mathbb{E}_{\tau \sim P_{\pi,T}}[R(\tau)]$$

environment
$$T(s_{t+1} \mid s_t, a_t)$$

o
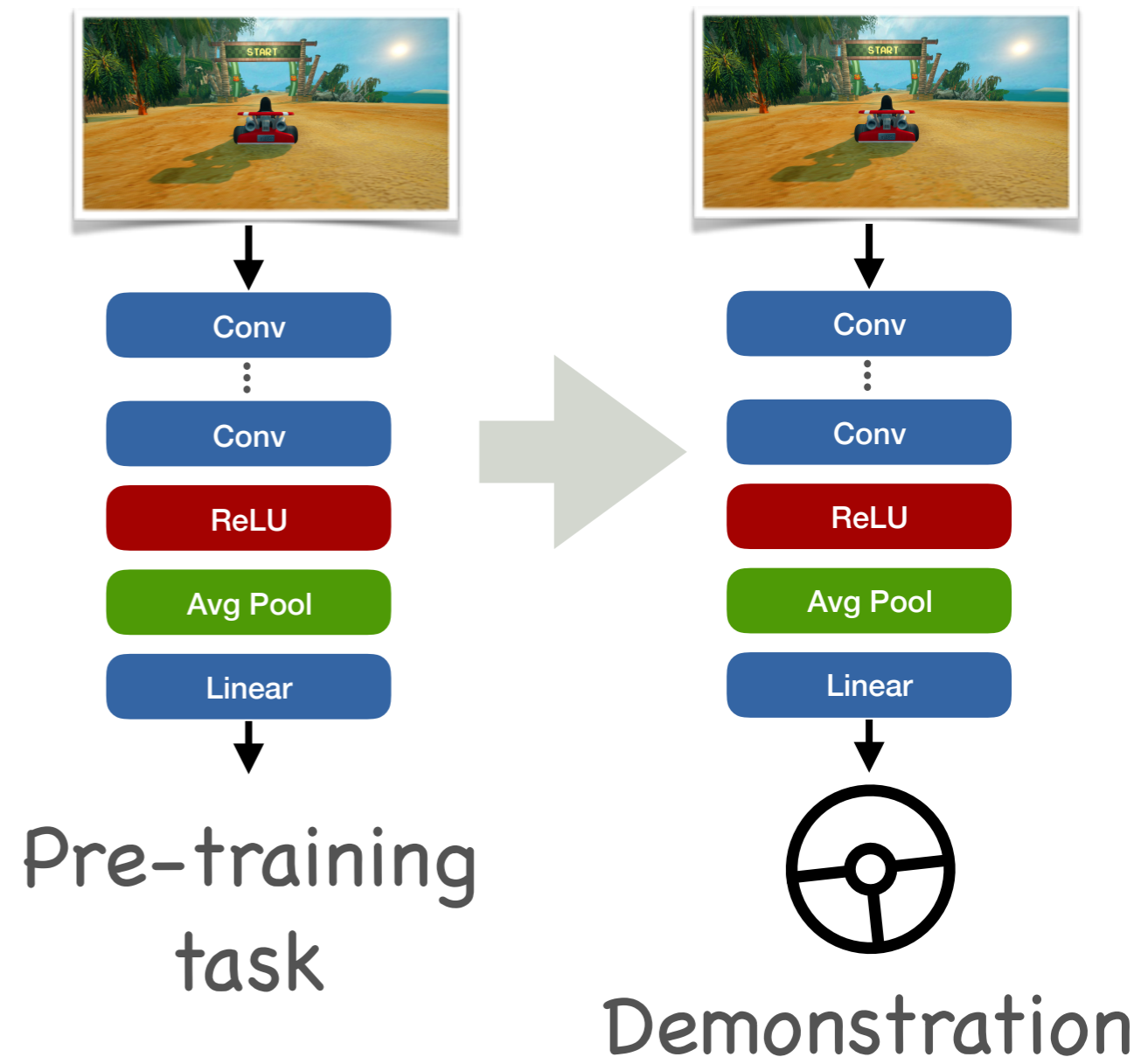
a

r_s

agent
$$\pi(a \mid o)$$

Imitation learning - Issues

- Expert annotations are sometimes expensive

- Drift from expert

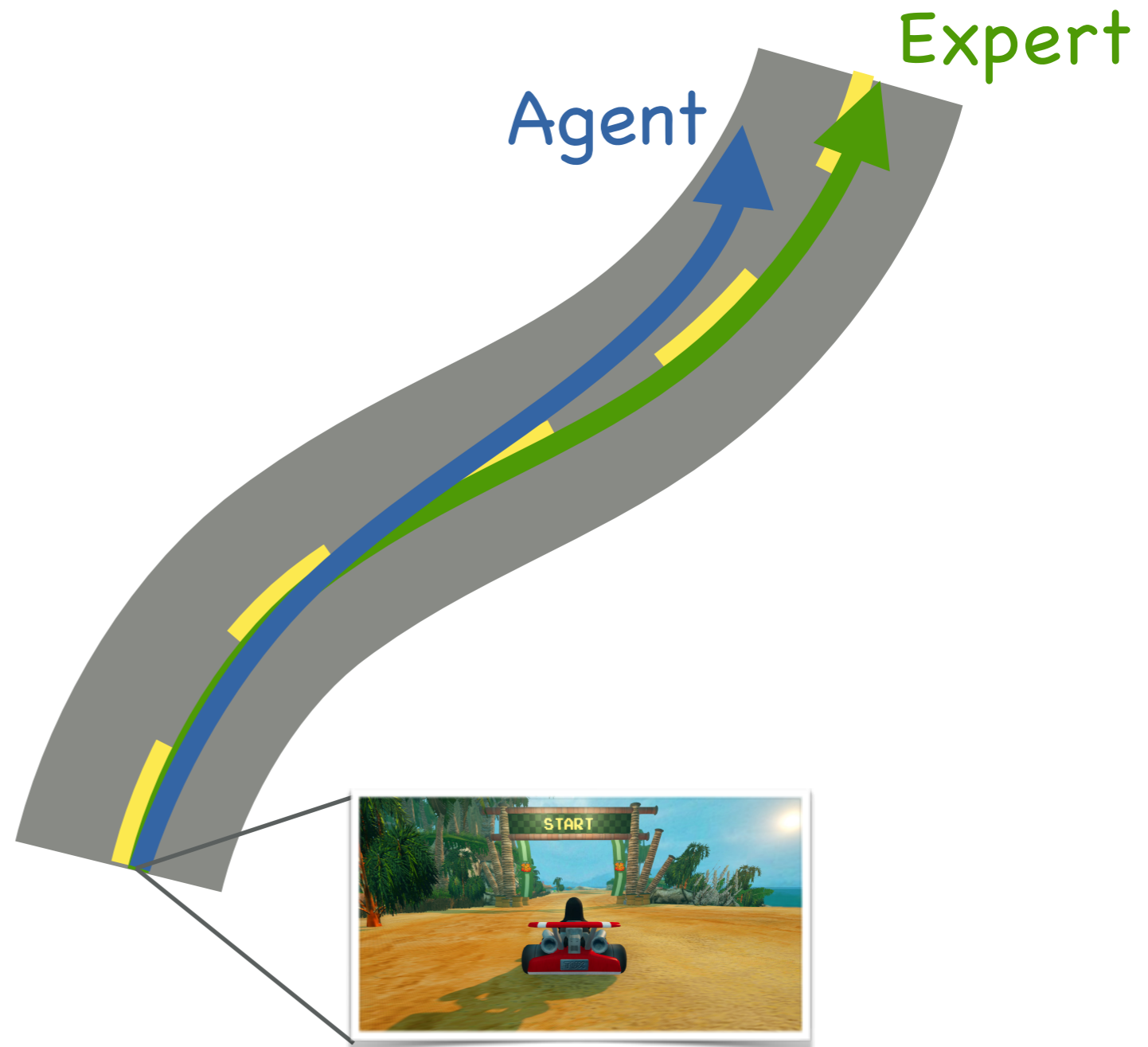 - Distribution mismatch between training and testing

Imitation learning - bag of tricks

- Use pre-trained architecture

- More robust to slight mismatch in observations

- Better generalization

Pre-training task

Demonstration

Imitation learning - bag of tricks

- Data augmentation

  - Encourage policy to get back on track

Imitation learning with tricks

- Easy to train

  - supervised learning

- Sometimes works

  - Major issues with drift / distribution mismatch

- Requires expert annotations

Exploring the Limitations of Behavior Cloning for Autonomous Driving, Codevilla etal. 2019