

# Handling Delay in Real-Time Reinforcement Learning

# What is Real-Time RL?

- Environment keeps changing while an agent infers its action.

# What is Real-Time RL?

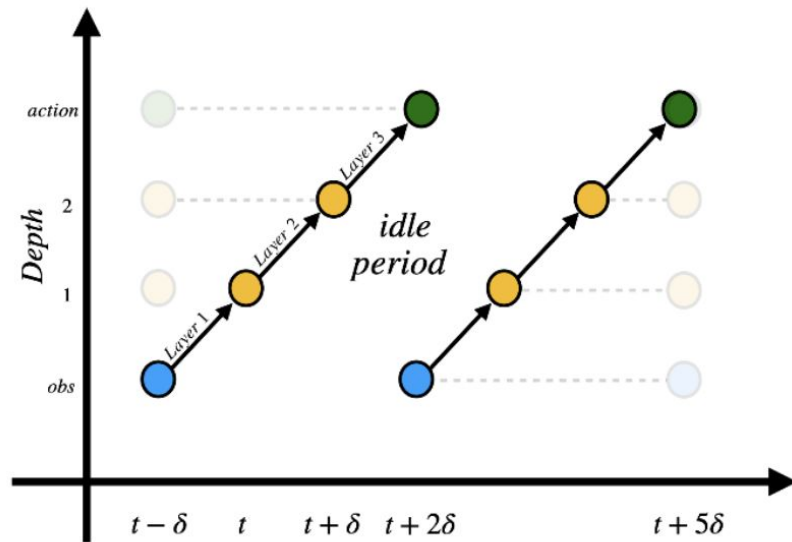
- Environment keeps changing while an agent infers its action.

Consequences:

- Observational delay.
- Inference time matters.

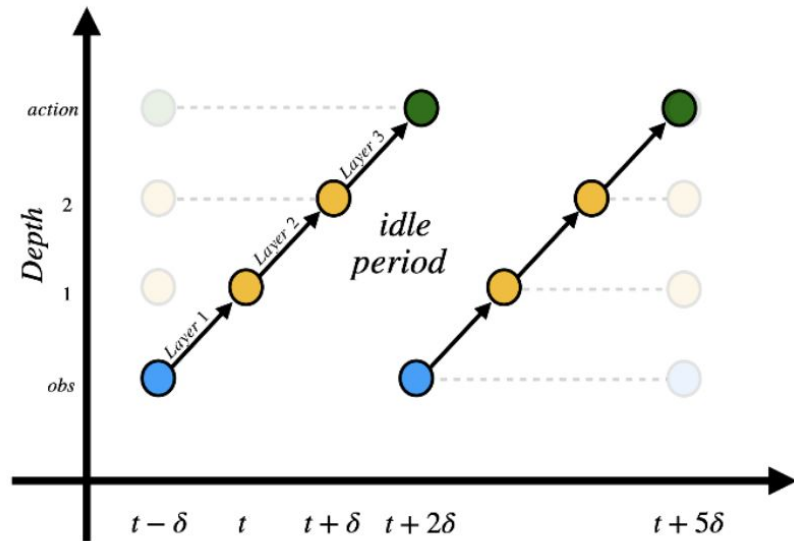
# Assumptions

- Each layer takes a fixed amount of time to execute.



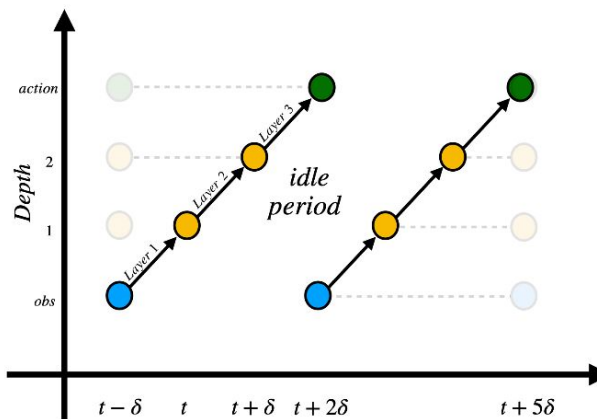
# Baseline setting

- Total delay time of  $3\delta$  per observation.
- There is a period of  $3\delta$  during which an agent cannot act.

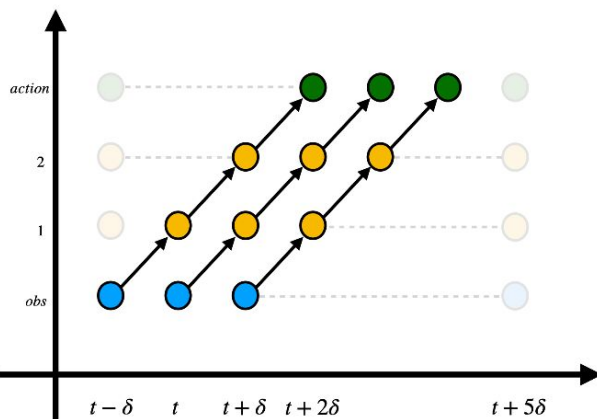


# How to improve over this baseline?

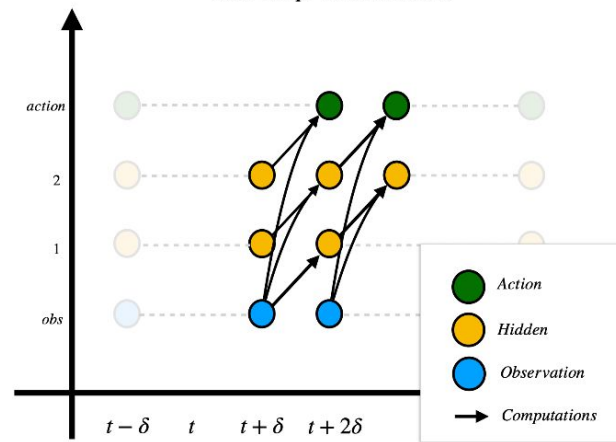
*Sequential Computations*



*Parallel Computations*



*With Skip Connections*

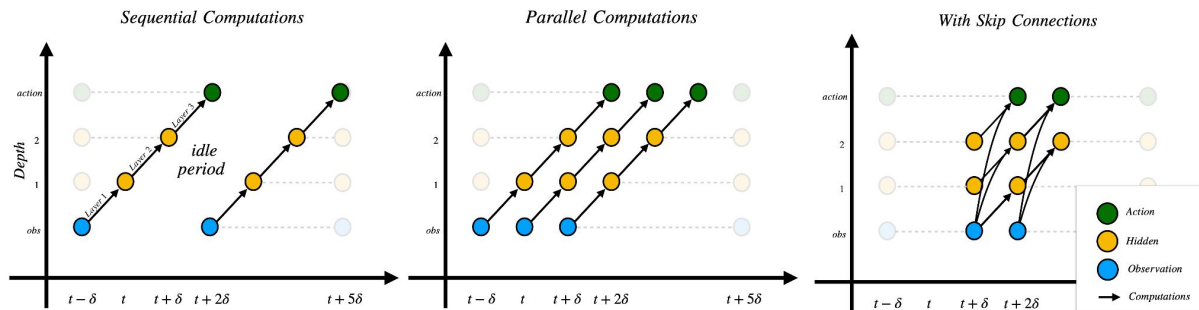


- Starting point

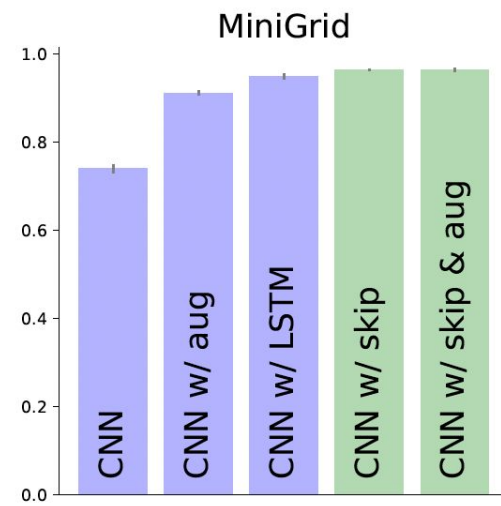
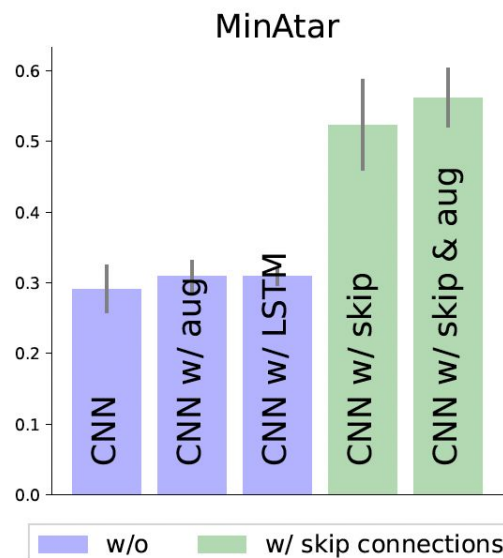
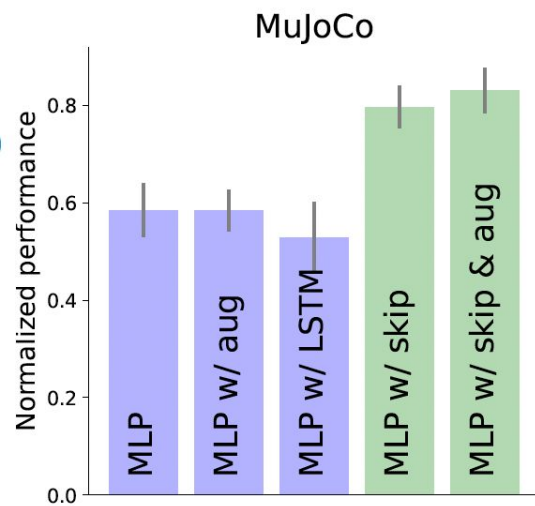
- Apply pipelining to address inaction

To address delay

- Apply temporal skip connection
- Augment a state with the last action



	Halfcheetah-v4	Walker2d-v4	Ant-v4	Hopper-v4
Sequential	0.246	0.651	0.516	0.456
Parallel	0.574	0.888	0.974	0.998
Parallel with skip connections	0.685	0.807	0.828	1.309





# Motivation

- Robotics
- On-device inference
- Neuromorphic computing.

Thank you for your attention