Imagination and Curiosity in Reinforcement Learning

Northwestern | McCORMICK SCHOOL OF ENGINEERING

Omkar Ranadive

What is Imagination?

- The ability to simulate ideas in our head
- We simulate these notions approximately based on our prior knowledge
- Imagination is useful in planning
- So in RL, the focus is to make agents plan better by “imagining”

Why would Imagination help?

- Two main types of RL - Model free, Model based
- Model free agents are difficult to train
- Difficult to form **perfect** models of the world in model based
- Imagination will help in forming **imperfect** models of the world
- Key idea: Combine model free agents with imagination

Model Architecture: Imagination Core

- Take the present observation and action and predict the next observation, reward
- Loop this process for multiple time steps - one trajectory
- Produce n such trajectories - Equivalent of thinking n scenarios

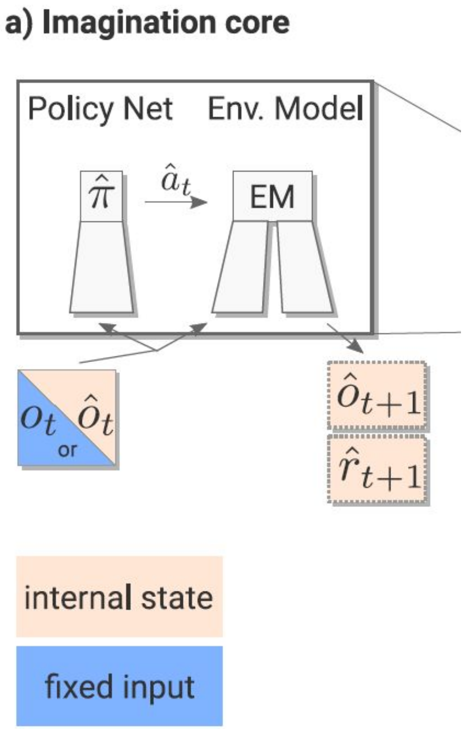
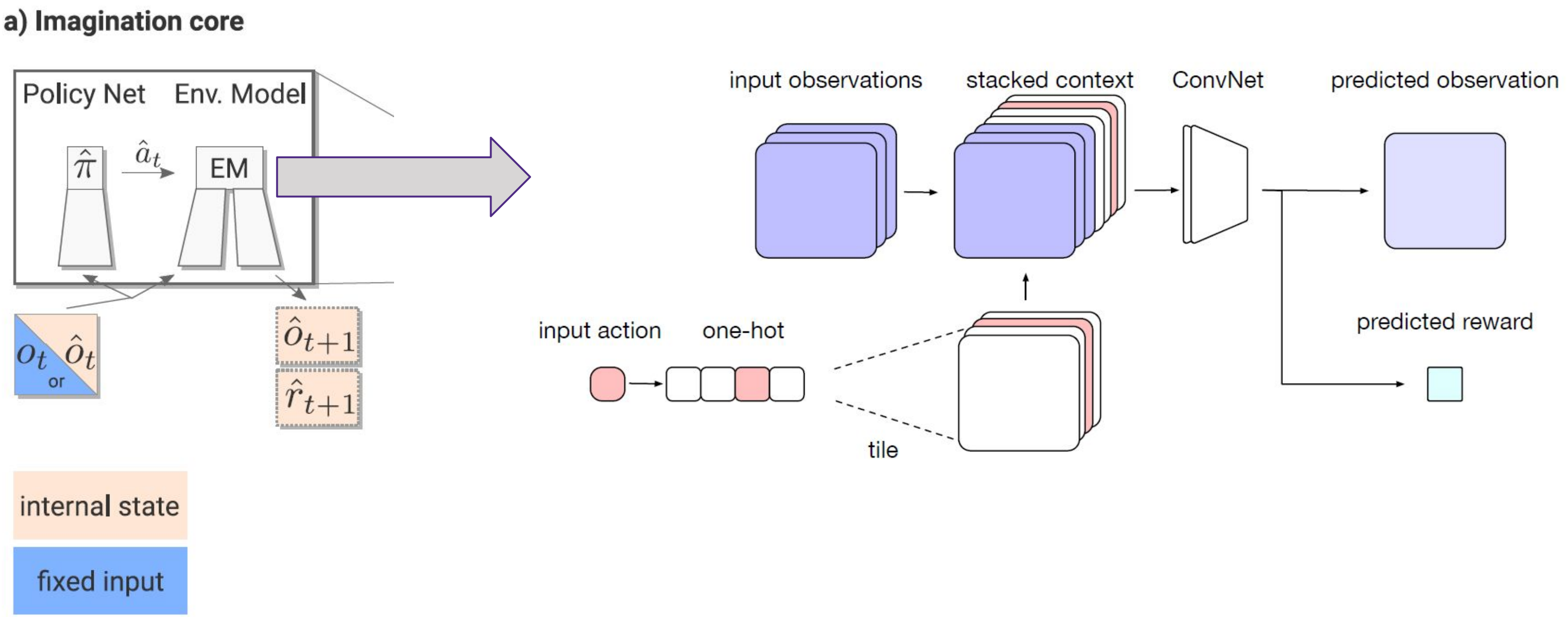


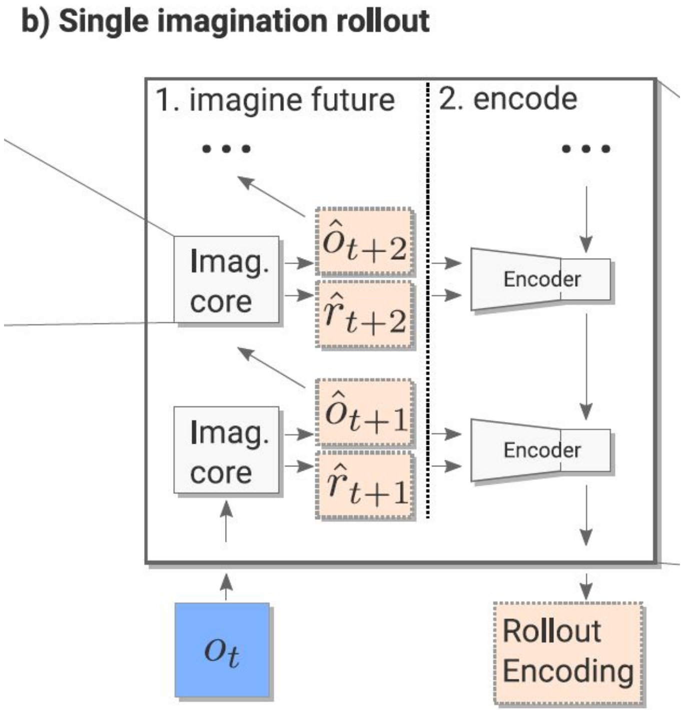
Image courtesy of Weber et al (2017)

Model Architecture: Environment Model



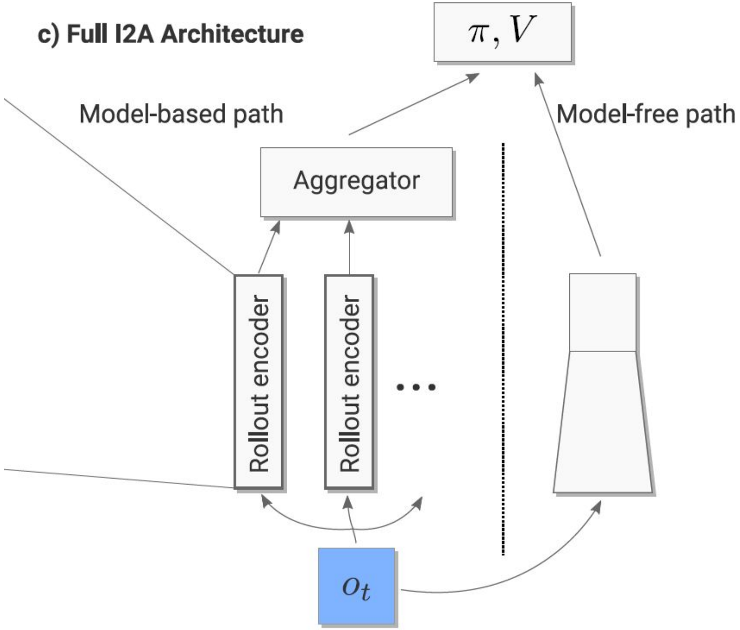
Model Architecture: Rollout Encoder

- The imagined trajectories are imperfect
- So, we want to discard information
- Idea - Only use imagined information (or part of it) if useful else discard
- Aggregate the encoded values

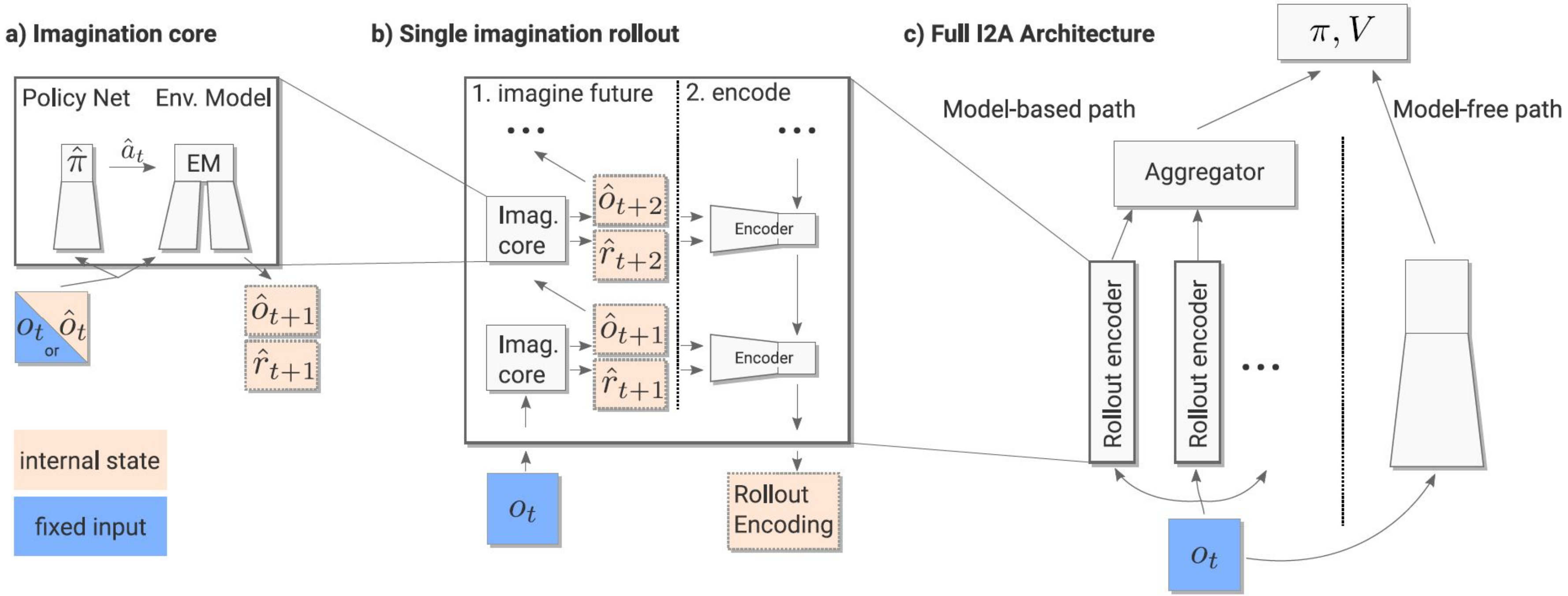


Model Architecture: Policy Module

- Take the aggregated information and combine it with model free output
- Without imagination, it's just a normal model free agent
- Hence, the title imagination "augmented" agents



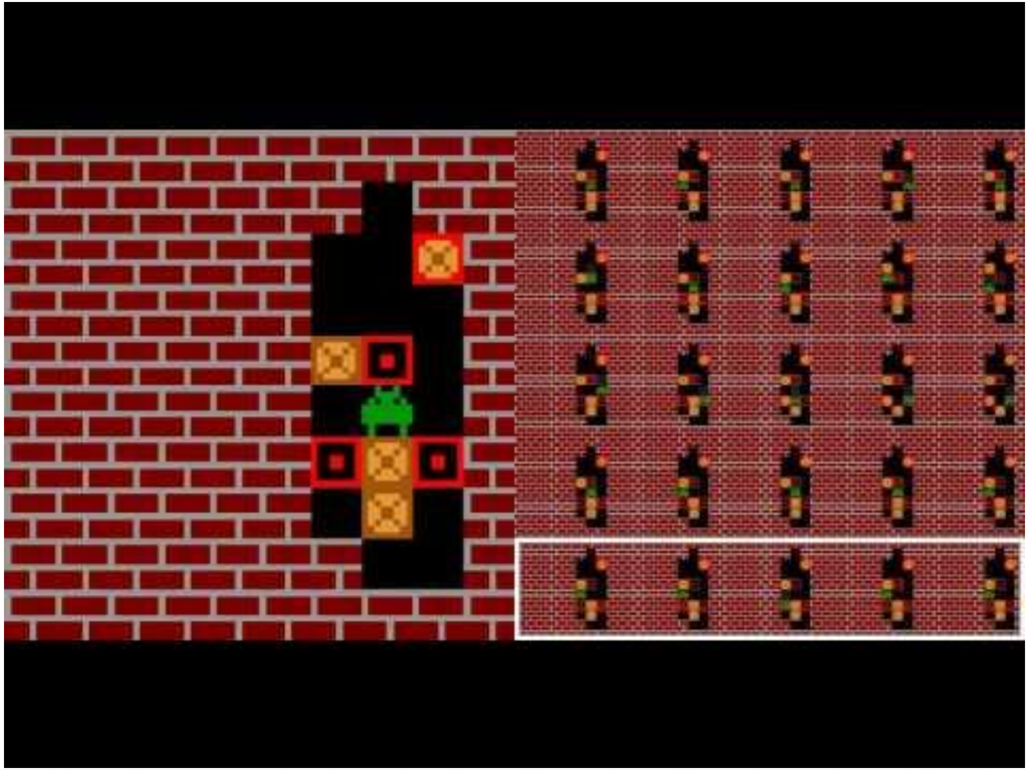
Model Architecture: Full View



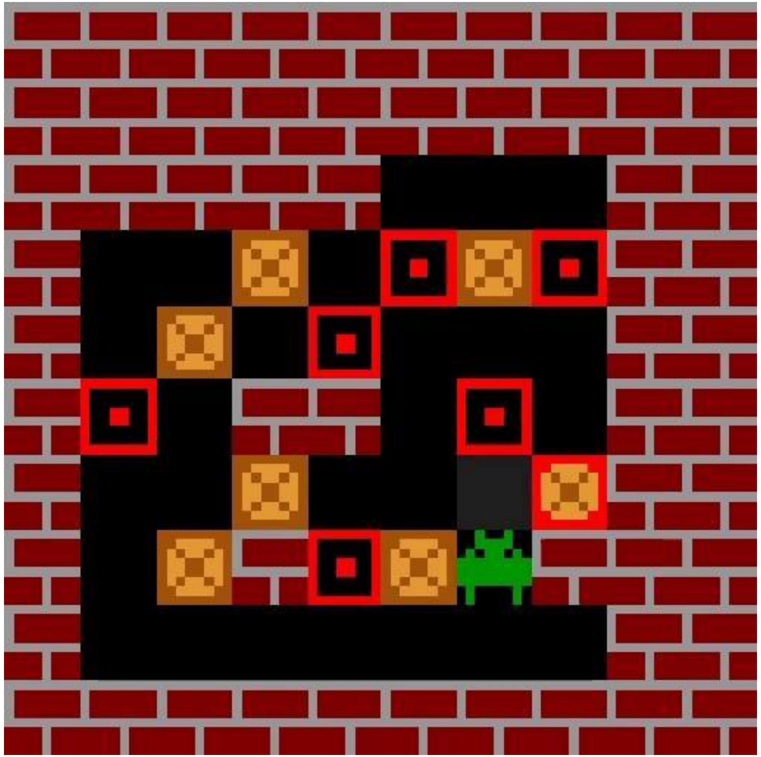
Training strategies

- The policy in environment model tries to imitate the model free policy
- Environment model was pre trained using a **partially trained** model free agent

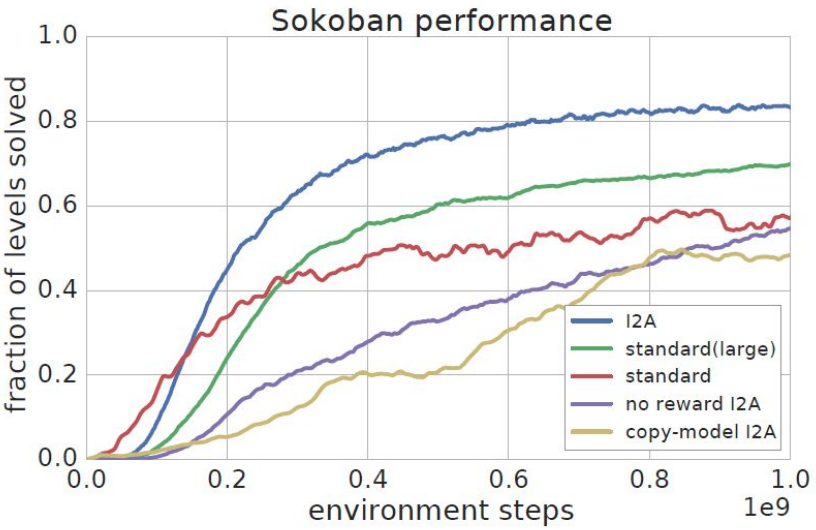
Results



Results: Generalization



Results: Comparison



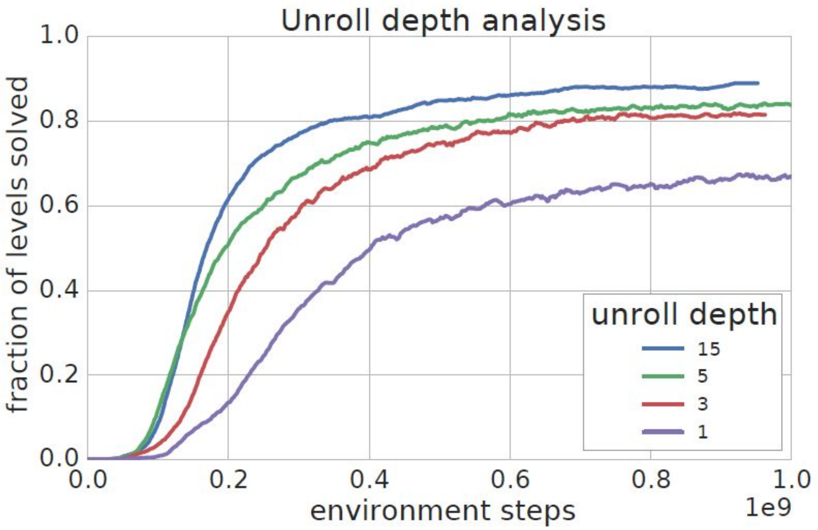


Image courtesy of Weber et al (2017)

What is Curiosity?

- Desire to learn something, seek new experiences
- Learn about things which we don't know much about
- In RL, this is useful for exploration
- So, make an agent explore the environment better by making it "curious"

Why would curiosity help?

- In RL, extrinsic rewards are usually sparse
- So, positive reinforcement happens only when we somehow encounter these rewards - difficult task
- Humans still explore the environment using motivation/curiosity
- Similarly, curiosity as an intrinsic reward would help the agent

Curiosity: Formal definition

- Curiosity is the error in predicting the consequence of its own actions
- Agent predicts the next state based on present state and action

$$p(\phi(x_{t+1})|x_t, a_t)$$

- The intrinsic reward is then: $r_t = -\log p(\phi(x_{t+1})|x_t, a_t)$
- Lower the probability higher the reward. So the agent gets rewards if it predicts hard to predict states

Curiosity: Inverse Dynamics Model

- Represent the high dimensional state using simpler representation
- Idea - Learn to predict action a_t based on s_t and s_{t+1}

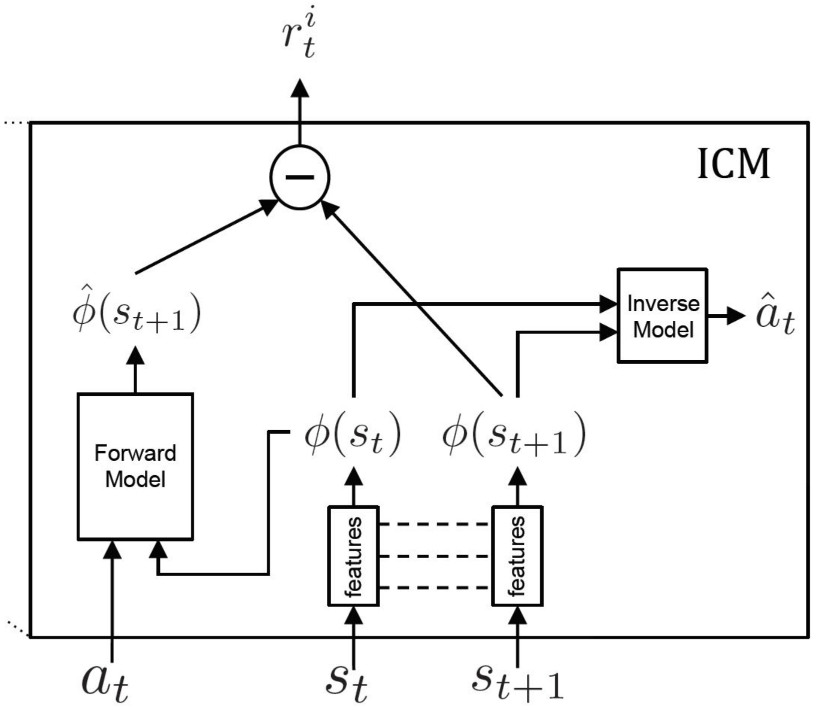
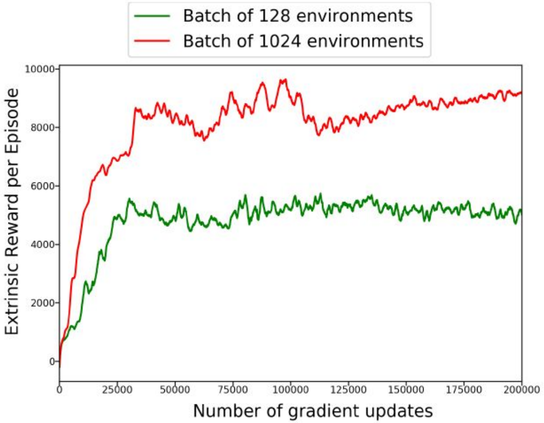


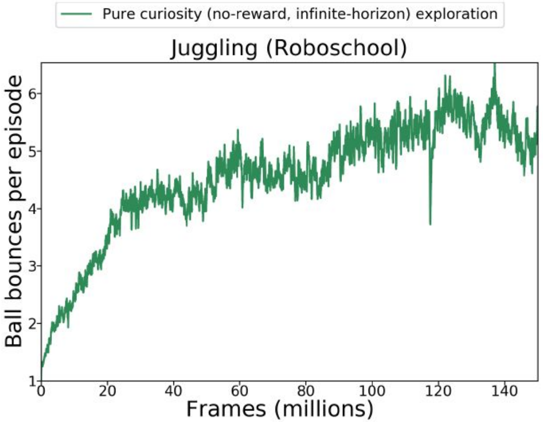
Image courtesy of Pathak et al (2017)

Results

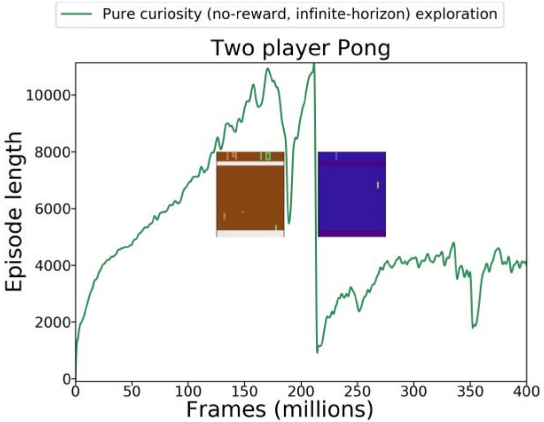
Results



(a) Mario w/ large batch



(b) Juggling (Roboschool)



(c) Two-player Pong

Image courtesy of Burda et al (2018)