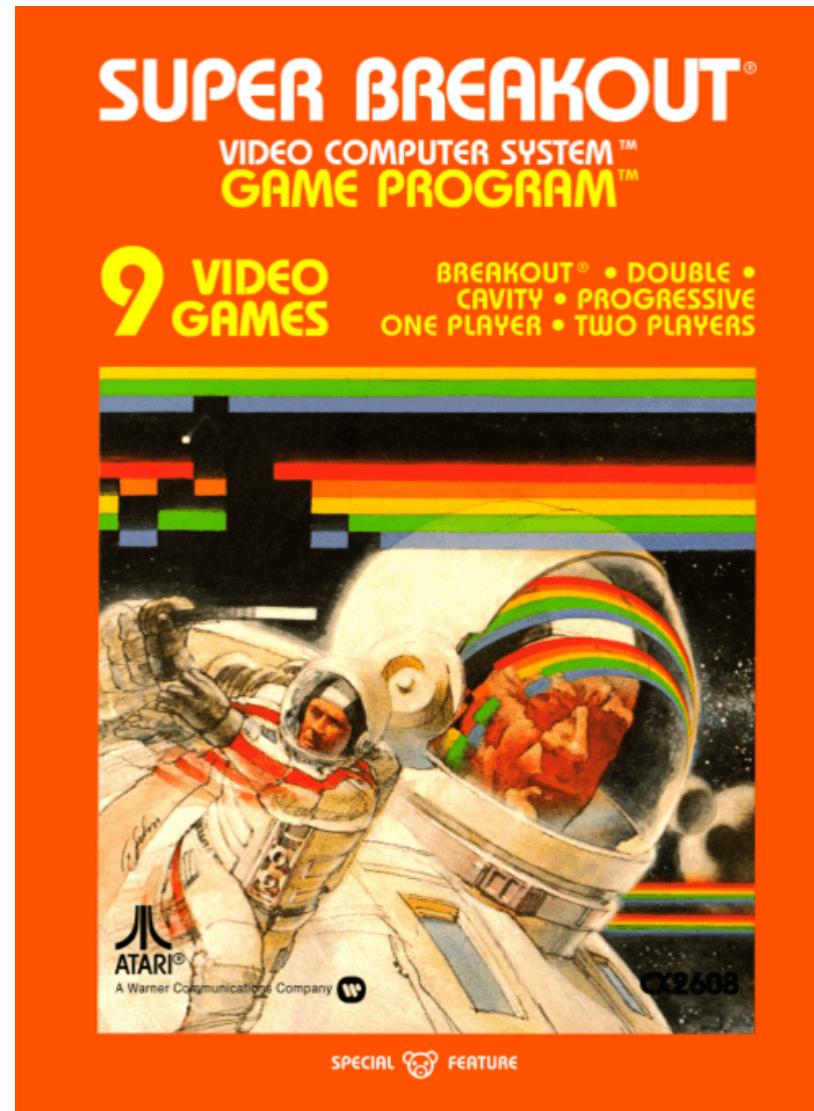


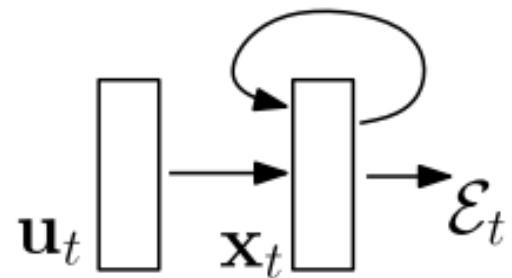
RL for POMDPs

4 Approaches

Approach 1: k-Markov



Approach 2: RNN (e.g. LSTM)

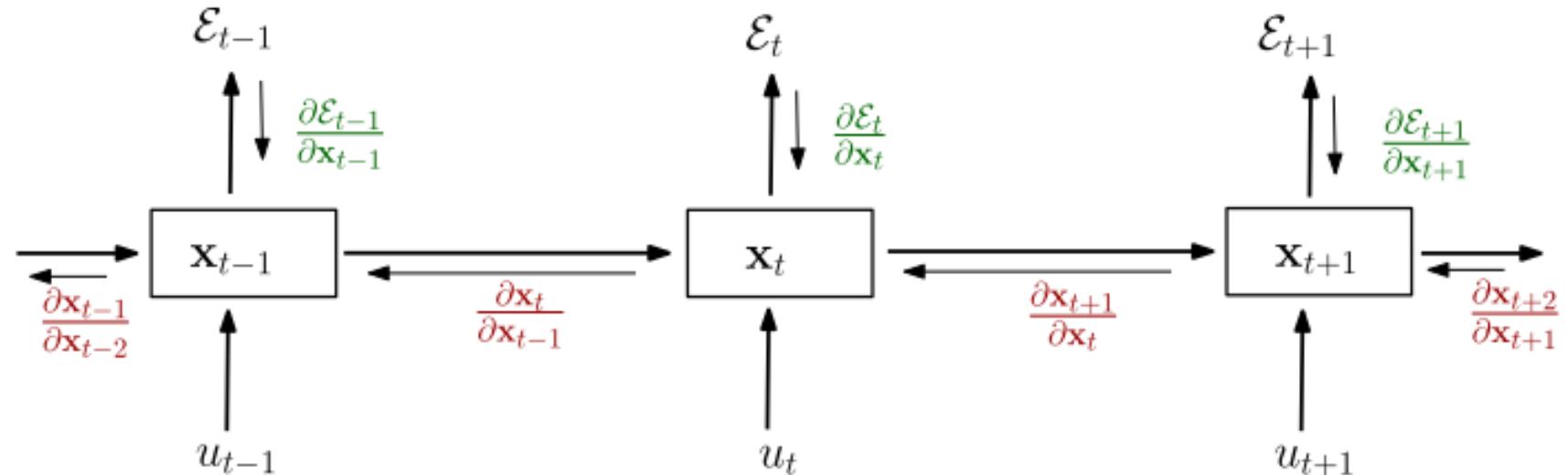


Input: u_t
State/Output: x_t

Cost: \mathcal{E}_t

$$\mathbf{x}_t = F(\mathbf{x}_{t-1}, \mathbf{u}_t, \theta)$$

$$\mathbf{x}_t = \mathbf{W}_{rec}\sigma(\mathbf{x}_{t-1}) + \mathbf{W}_{in}\mathbf{u}_t + \mathbf{b}$$



Approach 2: RNN (e.g. LSTM)

Input: x_t

Output: h_t

Cell state: c_t

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f)$$

Forget gate

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i)$$

Input Gate

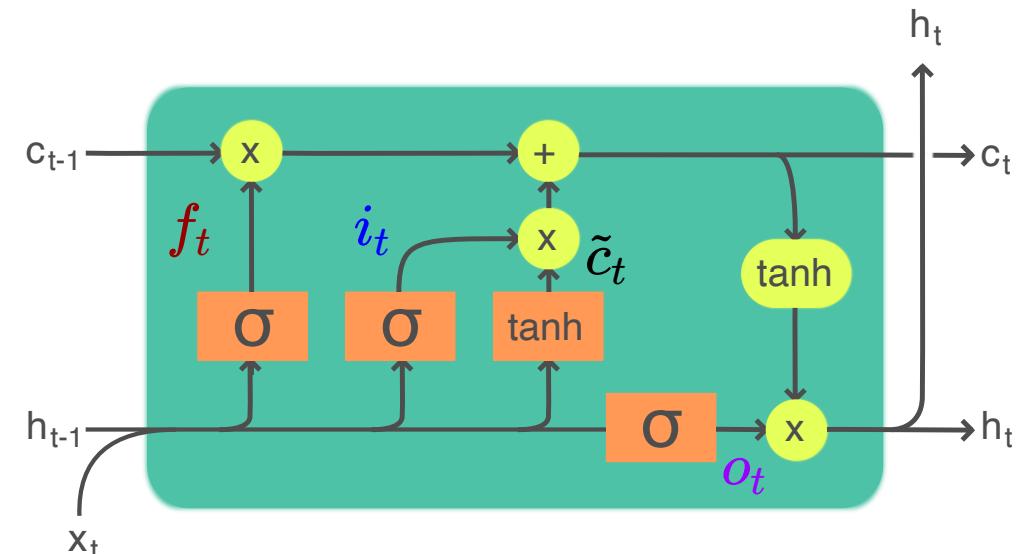
$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o)$$

Output Gate

$$\tilde{c}_t = \sigma_c(W_c x_t + U_c h_{t-1} + b_c)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$

$$h_t = o_t \odot \sigma_h(c_t)$$



Legend:

Layer

Componentwise

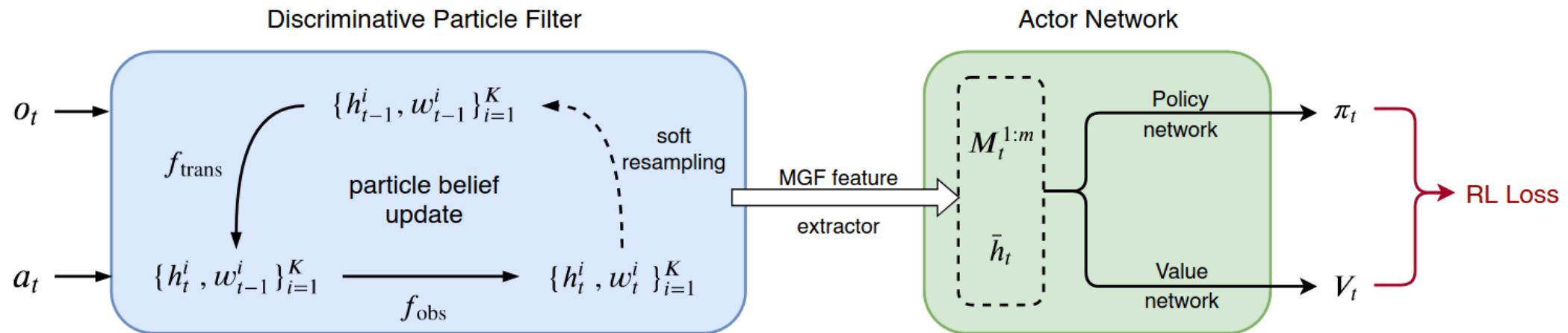
Copy

Concatenate



By Guillaume Chevalier - File:The_LSTM_Cell.svg, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=109362147>

Approach 3: Particle Filters and MGF Features



$$M_{\mathbf{X}}(\mathbf{v}) = \mathbb{E} \left[e^{\mathbf{v}^\top \mathbf{X}} \right]$$

$$M_j = \frac{d^j M_{\mathbf{X}}}{d\mathbf{v}^j} |_{\mathbf{v}=0}$$

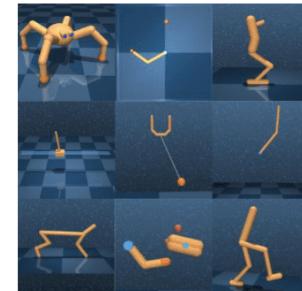
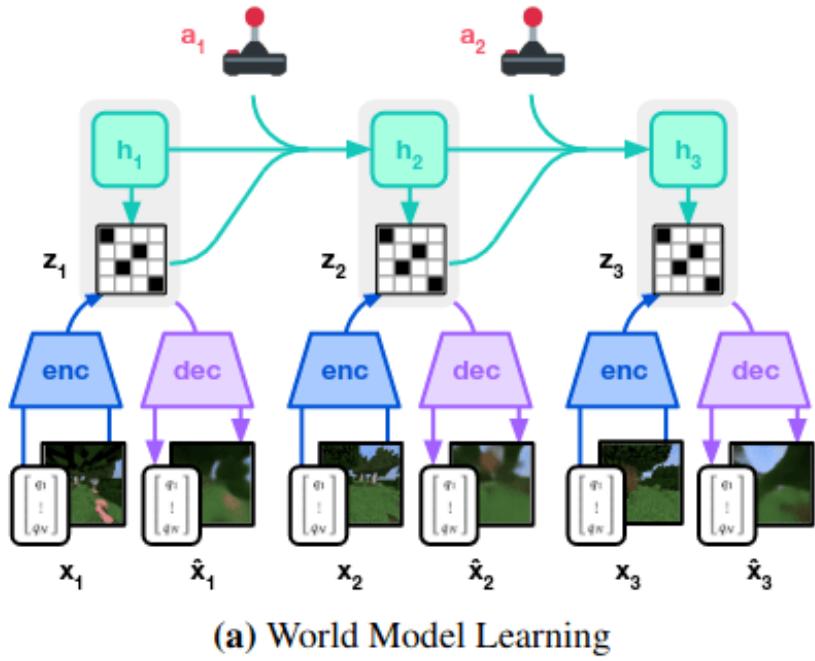
Published as a conference paper at ICLR 2020

DISCRIMINATIVE PARTICLE FILTER REINFORCEMENT
LEARNING FOR COMPLEX PARTIAL OBSERVATIONS

Xiao Ma¹, Peter Karkus¹, David Hsu¹ and Wee Sun Lee¹, Nan Ye²

¹National University of Singapore, ²The University of Queensland
 $\{\text{xiao-ma}, \text{karkus}, \text{dyhsu}, \text{leews}\}@comp.nus.edu.sg, \text{nan.ye@uq.edu.au}$

Approach 4: Latent Representations (e.g. World Models)



(a) Control Suite



(b) Atari



(c) DMLab



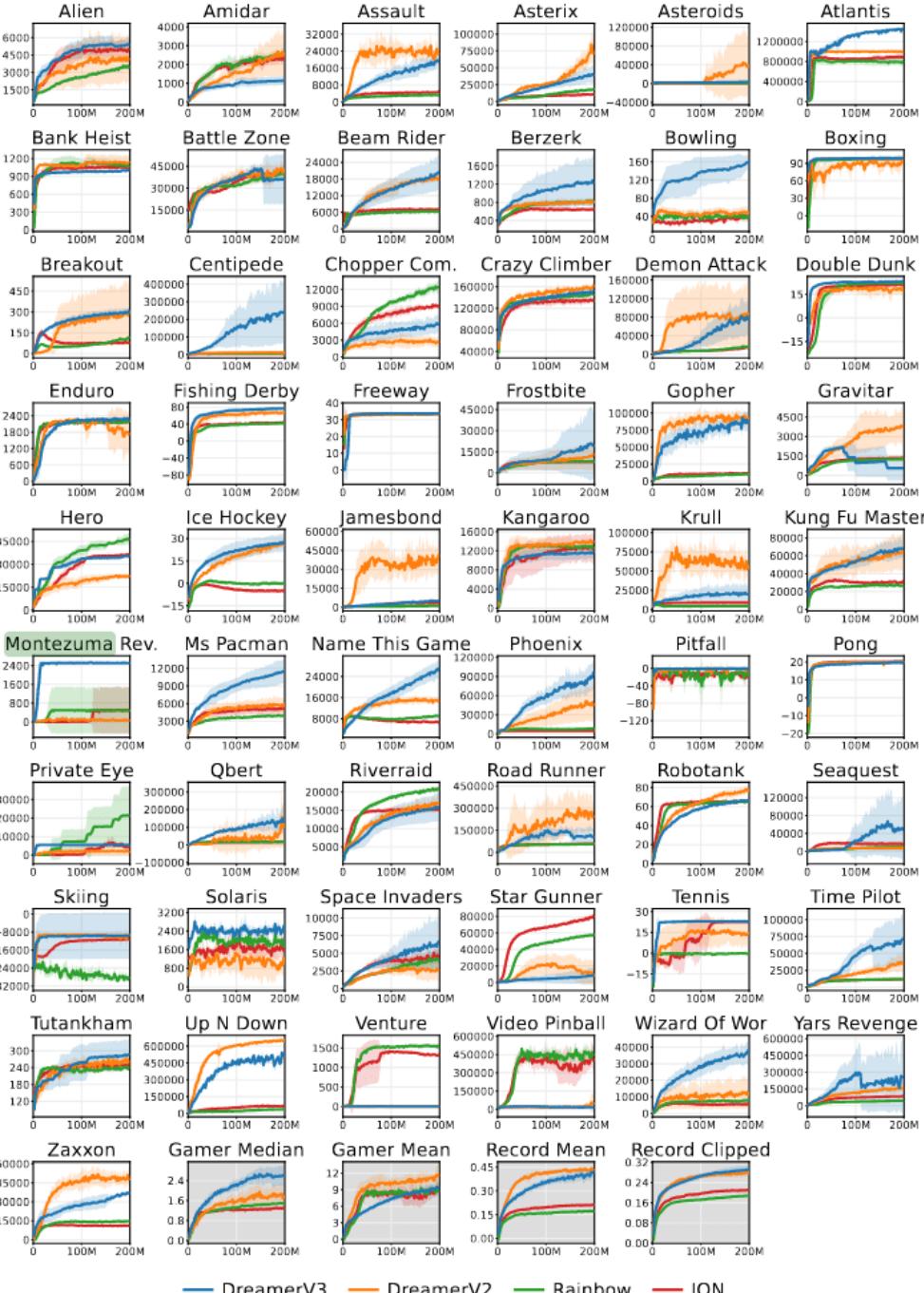
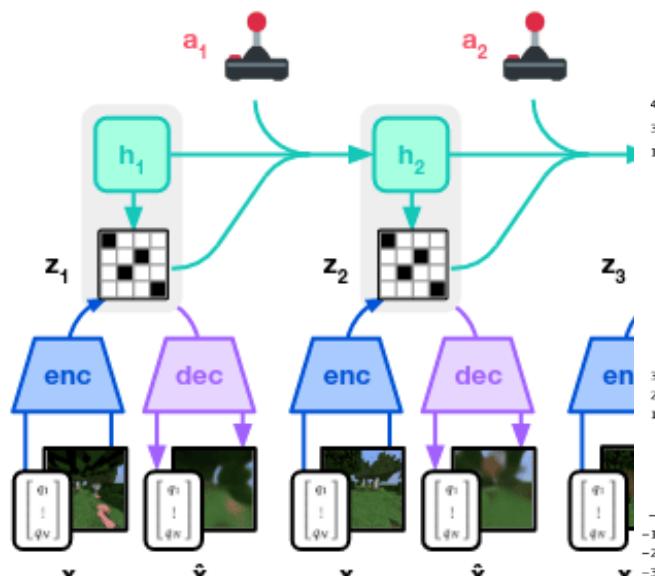
(d) Minecraft

Mastering Diverse Domains through World Models

Danijar Hafner,^{1,2} Jurgis Pasukonis,¹ Jimmy Ba² Timothy Lillicrap¹

¹DeepMind ²University of Toronto

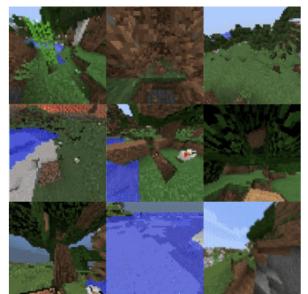
Approach



Presentations



(c) DMLab



(d) Minecraft

g Diverse Domains through World Models

Hafner,^{1,2} Jurgis Pasukonis,¹ Jimmy Ba² Timothy Lillicrap¹

¹DeepMind ²University of Toronto