Locality Sensitive Sparse Encoding for Learning World Models Online

Zichen Liu[12], Chao Du[1], Wee Sun Lee[2], Min Lin[1]

[1]Sea AI Lab [2]NUS

ICLR 2024

MBRL needs **online** world models

- Tackling non-stationarity, forgetting…
- **F**ollow-**T**he-**L**eader Online Learning

$$\forall t, \mathbf{w}_t = \arg\min_{\mathbf{w}\in\mathcal{S}} \sum_{i=1}^{t-1} \ell_i(\mathbf{w})$$

- Deep world models with replay?

$$\forall t, \theta_t = \arg\min_{\theta} \mathbb{E}_{\mathcal{D}_{t-1}} \|f_\theta(\mathbf{s}, \mathbf{a}) - \mathbf{s}'\|_2^2$$

Linear models learn online analytically & incrementally

$$\forall t, \mathbf{W}_t = \arg\min_{\mathbf{W}\in\mathbb{R}^{D\times S}} \|\mathbf{\Phi}_{t-1}\mathbf{W} - \mathbf{Y}_{t-1}\|_F^2$$

$$\Leftrightarrow$$

$$\mathbf{W}_{t+1} = \mathbf{A}_t^{-1}\mathbf{B}_t, \text{where } \mathbf{A}_t = \mathbf{\Phi}_t^\top \mathbf{\Phi}_t, \mathbf{B}_t = \mathbf{\Phi}_t^\top \mathbf{Y}_t$$

Online model learning

$t = 0, \mathbf{A}_0 \in \mathbb{R}^{D\times D} = \mathbf{0}, \mathbf{B}_0 \in \mathbb{R}^{D\times S} = \mathbf{0}$

while true:

$t \leftarrow t + 1$

$\mathbf{A}_t \leftarrow \mathbf{A}_{t-1} + \phi(\mathbf{x}_t)\phi(\mathbf{x}_t)^\top$

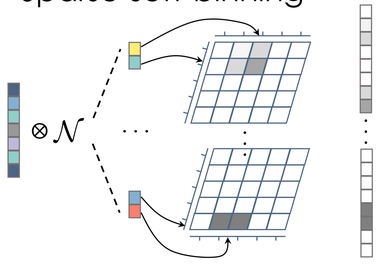
$\mathbf{B}_t \leftarrow \mathbf{B}_{t-1} + \phi(\mathbf{x}_t)\mathbf{y}_t^\top$

$\mathbf{W}_{t+1} \leftarrow \mathbf{A}_t^{-1}\mathbf{B}_t$

High-d random non-linear features boost capacity while **sparsity** bounds computation

Locality **S**ensitive **S**parse **E**ncoding

- Random projection to high-d
- Sparse soft binning



Sparse online model learning

$t = 0, \mathbf{A}_0 \in \mathbb{R}^{D\times D} = \mathbf{0}, \mathbf{B}_0 \in \mathbb{R}^{D\times S} = \mathbf{0}$

while true:

$t \leftarrow t + 1$

$s \leftarrow \text{nonzero\_index}(\phi(\mathbf{x}_t))$

$\mathbf{A}_{t,ss} \leftarrow \mathbf{A}_{t-1,ss} + \phi_s(\mathbf{x}_t)\phi_s(\mathbf{x}_t)^\top$

$\mathbf{B}_{t,s} \leftarrow \mathbf{B}_{t-1,s} + \phi_s(\mathbf{x}_t)\mathbf{y}_t^\top$

$\mathbf{W}_{t+1,s} \leftarrow \mathbf{A}_{t,ss}^{-1}(\mathbf{B}_{t,s} - \mathbf{A}_{t,s\bar{s}}\mathbf{W}_{t,\bar{s}})$

Empirical results

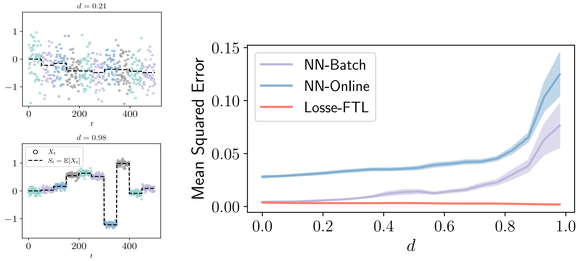
- Supervised learning experiments

Representation power

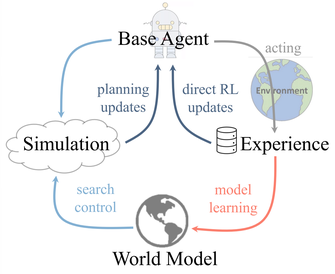
	Sparse	Real-value	9	16	25	36	49
NN	-	-	0.36	0.35	0.34	0.34	0.33
Fourier	✗	✓	0.32	0.39	0.67	1.00	1.40
ReLU	✓	✓	0.35	0.34	0.37	0.39	0.43
Tile Code	✓	✗	0.36	0.39	0.51	0.61	0.73
Losse	✓	✓	**0.28**	**0.29**	**0.31**	**0.34**	**0.40**

Error of MNIST reconstruction

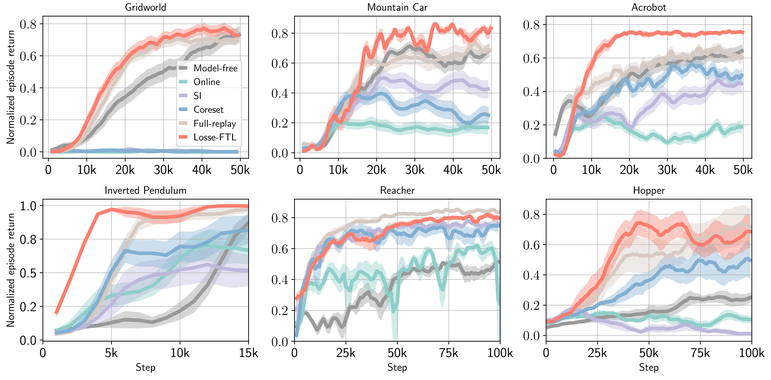
Non-stationary stream learning



- Reinforcement Learning experiments



Dyna framework



- Learns incrementally online, without replay buffer
- Outperforms several continual learning techniques