Context and Diversity Matter: Emergence of In-Context Learning in World Models

Fan Wang*, Zhiyuan Chen*, Yuxuan Zhong*, Sunjian Zheng, Pengtao Shao,
Bo Yu, Shaoshan Liu, Jianan Wang, Ning Ding, Yang Cao , Yu Kang

SHENZHEN INSTITUTE of ARTIFICIAL INTELLIGENCE AND ROBOTICS for SOCIETY
深圳市人工智能与机器人研究院

中国科学技术大学
University of Science and Technology of China

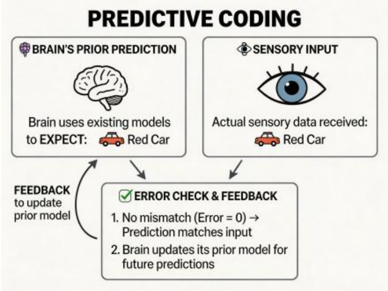
Astribot

Corresponding to: fanwang.px@gmail.com, shaoshanliu@cuhk.edu.cn

Motivations and Introductions

Enable Self-Adaptation World Models

- Existing world models: incapable of adapting in non-stationary / long-tail scenarios
- Human beings and animals learn and adapt on-the-fly through *Predictive Coding*
- In-Context Learning (ICL): widely studied in LLMs, but underexplored in world models



Research Questions

- **Can world models perform ICL, and what factors influence this capability?**

Contributions

- **Theoretical analyses** uncover characteristics of ICL of world model in MDP and POMDP
- **Empirical Validation:**

  Propose long-context linear-attention world model (L2World)

  Consistent with the theoretical analyses

Theoretical Analyses (I)

Two modes of ICL in World Models: Environment Recognition & Environment Leanring

Definition： World Models

$$\textbf{World Model: } \hat{o}_{t+1} \sim \hat{p}_\theta(\cdot|q_t) = f_\theta(q_t), \text{with } q_t = (s_t, a_t) \text{ (MDP)}$$
$$\text{or } q_t = (o_{t-\Delta t}, a_{t-\Delta t}, \ldots, o_t, a_t) \text{ (POMDP)} .$$

Definition： In-Context Learning

$$\forall T_1 > T_2, D[\hat{p}_\theta(\cdot|q_t, C_{T_1})||p_e(\cdot|q_t)] < D[\hat{p}_\theta(\cdot|q_t, C_{T_2})||p_e(\cdot|q_t)],$$

Environment Recognition (ER)

- **Assume a combination of seen (training) environments**
- **Dependent on both context (for environment recognition) & parametric memory (environment-specific inference)**

$$\hat{p}_{\theta,ER}(o_{t+1}|q_t, C_T) = \sum_{e \in \mathcal{E}} \underbrace{\hat{p}_\theta(e|q_t, C_T)}_{\substack{\text{Environment}\\ \text{Recognition}}} \cdot \underbrace{\hat{p}_{\theta,e}(o_{t+1}|q_t)}_{\substack{\text{Environment-Specific}\\ \text{World Model}}}$$

Environment Learning (EL)

- **Fully context dependent inferences**

$$\hat{p}_{\theta,EL}(o_{t+1}|q_t, C_T) = \frac{p(q_t, o_{t+1}|C_T)}{p(q_t|C_T)}$$

Theoretical Analyses (II)

Acquire insights through analyzing error upper bounds

Number of Seen Environments

$$TV(\hat{p}_{ER}, p_{e_0}) \leq \min[\underbrace{\alpha/3 \cdot (|\mathcal{E}| - 1) \cdot T^{-1/2}}_{Recognition\ Error}, \underbrace{\max_{e_1, e_2 \in \mathcal{E}} TV(p_{e_1}, p_{e_2})}_{Diversity}] + \underbrace{\min_{e \in \mathcal{E}} TV(\hat{p}_{\theta,e}, p_{e_0})}_{Best\ Matching\ Error}$$

$$TV(\hat{p}_{EL}, p_{e_0}) \leq \underbrace{\sqrt{2|O||S||A| log(4|O|/\delta)}}_{Environment\ Complexity} \cdot T^{-1/2}$$

Context lengths

with probability $1 - \delta$, *and* $T > 4|S|^2|A|^2 \log(4|S||A|/\delta)$

Assumptions: model "choose" ICL mode by lower error upper bounds

Insights

- **Lower environmental complexity and a greater number of environments favor EL over ER**
- **Long context and environment diversity are key to both ER and EL**
- **Over-training and powerful IWL facilitate ER over EL**

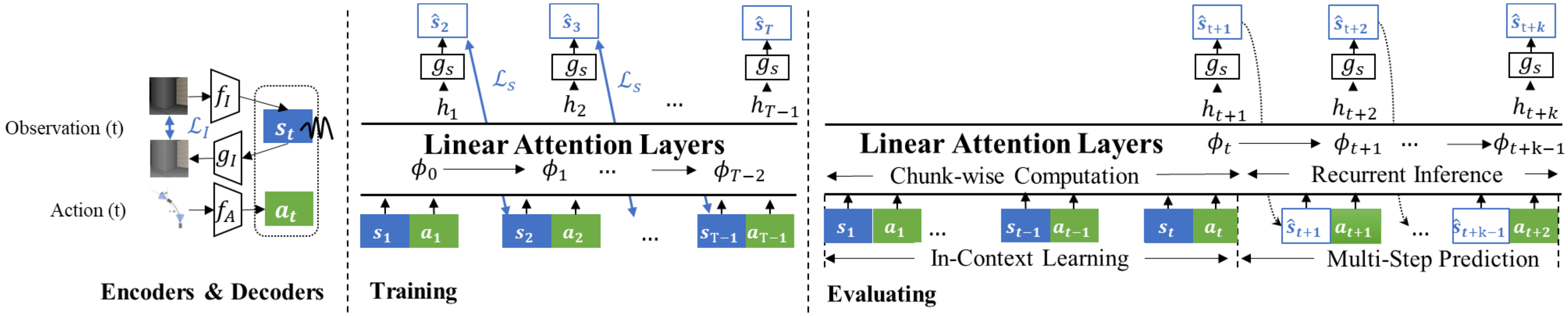
Principles to Maximize IC-EL capability

- **Long and consistent trajectories**
- **Sufficiently diverse trajectories (relative to the environment complexity)**
- **Long-context memory-efficient model structures**

Empirical Validation (I)

Long-context linear-attention world model (L2World)

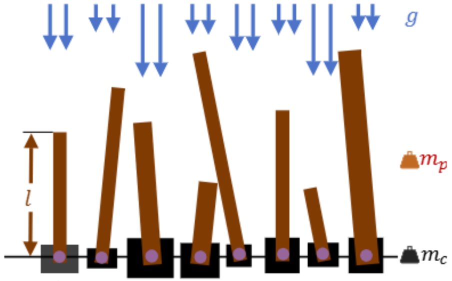
- Each image / action preprocessed by VAE / MLP encoder to acquire one single token
- Linear Attention Layers (Gated Slot Attention)
- Context window up to **10K images**
  - Dreamer-V3 (Hafner et al., 2025): LSTM, context window <1K
  - NWM (Bar et al., 2025) : Diffusion, context window = 4



Empirical Validation (II)

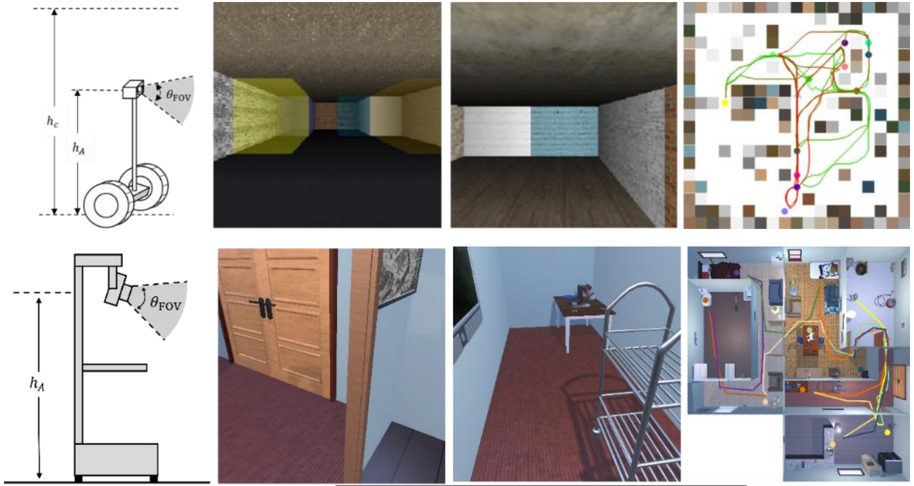
Experiment Settings

- Random cartpoles
- Random houses navigation (Mazes, ProcTHOR (Deitke et al., 2022))



	g	m_c	m_p	l
Scope 1	[8,12]	[0.8, 1.2]	[0.08, 0.12]	[0.4, 0.6]
Scope 1+2	[2,16]	[0.5, 2.0]	[0.05, 0.20]	[0.20, 1.0]

Random cartpoles

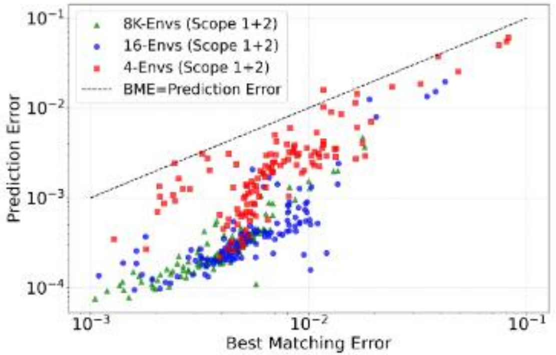


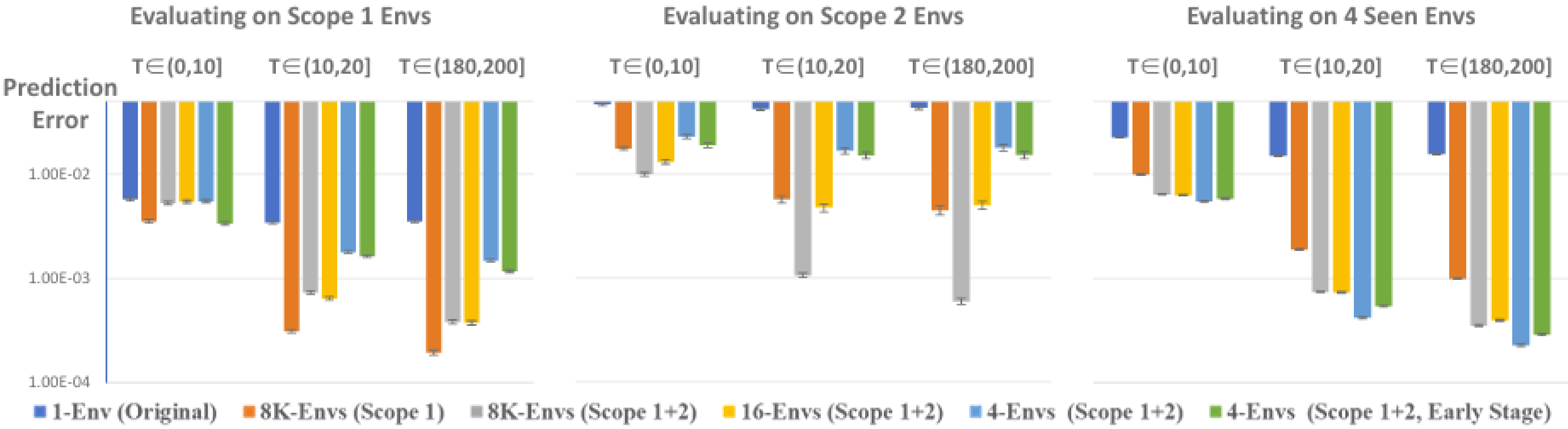
Random houses navigation

Empirical Validation (III)

Results – Random CartPoles

- BME indeed controls the upper error bound
- Diversity and number of environments matters for unseen envs





Empirical Validation (III)

Results – Navigation World Models

- Pre-training in Mazes (assets free)
- Post-training in ProcTHOR
- Identical data scales, more diversity and longer trajectories means better generalization (to unseen environments)

Model	Pre-train	Post-train	T=1	T=10	T=100	T=1000	T=10000
L2World	-	ProcTHOR-5K	15.49	18.22	19.02	19.74	19.81
L2World	Maze-32K-L	ProcTHOR-5K	16.46	**20.23**	**21.05**	**21.89**	**22.04**
L2World	Maze-32K-S	ProcTHOR-5K	**19.80**	19.45	19.86	20.57	20.61
L2World	Maze-128-L	ProcTHOR-5K	19.16	19.60	20.20	20.94	16.46
Dreamer	-	ProcTHOR-40K	19.82	22.61	23.99	23.51	22.76
NWM	-	ProcTHOR-40K	18.30	21.41	21.11	21.02	20.08
L2World	-	ProcTHOR-40K	**21.57**	22.67	23.39	24.92	22.98
L2World	Maze-32K-L	ProcTHOR-40K	17.21	**22.81**	**24.32**	**25.40**	**23.94**

Model	Seen					Unseen				
	T=1	T=10	T=100	T=1000	T=10000	T=1	T=10	T=100	T=1000	T=10000
L2World (Maze-32K-L)	16.80	20.97	23.11	24.65	25.05	16.37	**21.24**	**23.17**	**24.66**	**24.65**
L2World (Maze-32K-S)	18.57	19.28	19.67	20.21	20.48	**18.45**	19.24	19.63	20.29	20.31
L2World (Maze-128-S)	19.47	20.39	20.58	22.02	21.77	18.01	18.63	19.00	19.67	19.63
L2World (Maze-128-L)	18.54	20.86	**23.32**	**25.65**	**26.00**	17.54	19.43	20.96	21.54	21.52
Dreamer (Maze-32K-L)	16.40	**21.82**	19.24	21.26	21.89	16.81	20.48	21.40	22.65	22.12
Dreamer (Maze-128-L)	17.13	20.64	21.83	22.20	22.43	14.26	14.54	14.09	13.46	13.50
NWM (Maze-32K-L)	**20.84**	20.21	19.19	22.32	21.06	16.20	16.71	17.00	17.37	17.85

Additional Results

AR 1: Empirical Evidences of EL vs ER

- EL is more context-dependent:

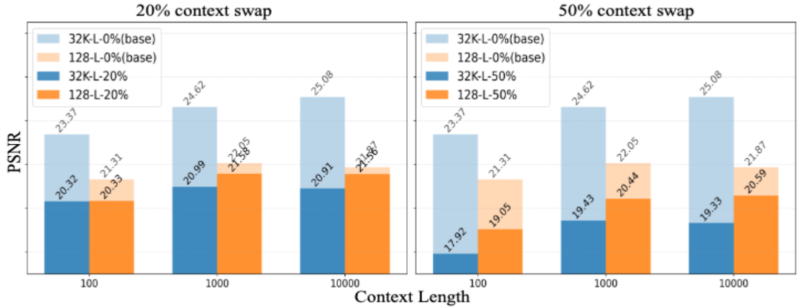
  more sensitive to context disturbances

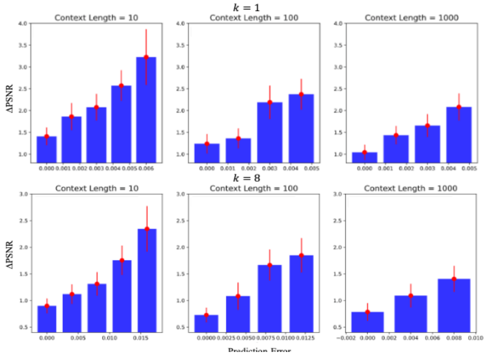
AR 2: ICL of world models aligns well with predictive coding

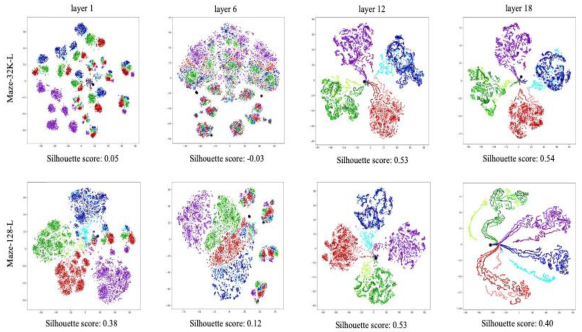
- The larger the current prediction error,

  the more the model gain in the next step

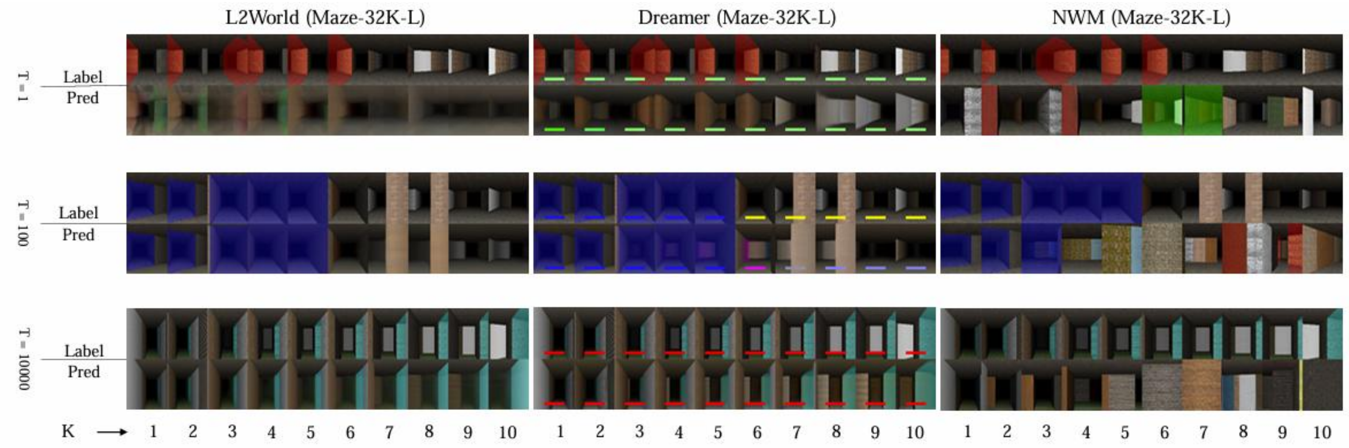
AR 3: Memory states in Linear attention effectively captures global mapping through ICL only







Thanks!



A demo cases of the self-adapation of world models through ICL (Comparison among L2World, Dreamer, NWM)

Codes