AI DAY 2020 RISING TO THE CHALLENGES

Predictive Coding for Locally Linear Control

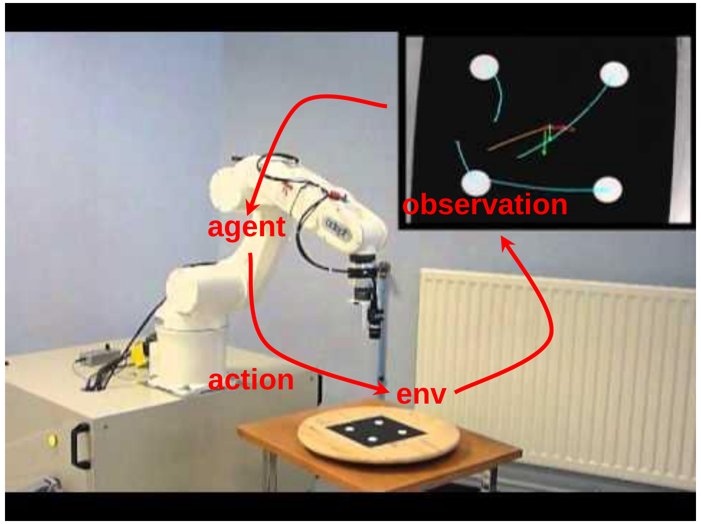
Rui Shu*, Tung Nguyen*, Yinlam Chow, Tuan Pham, Khoat Than, Mohammad Ghavamzadeh, Stefano Ermon, Hung Bui

SPEAKER:
Tung Nguyen – VinAI Resident

VinAI RESEARCH

Motivation

Decision making from high-dimensional observations

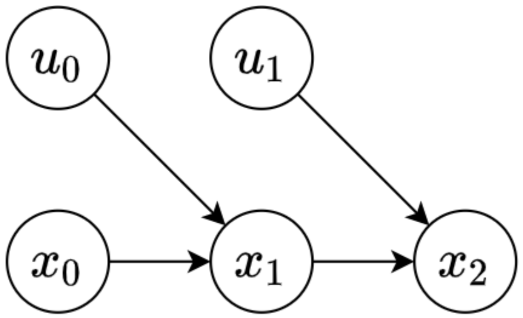


Example: Robot arm manipulation from visual inputs

Problem Formulation

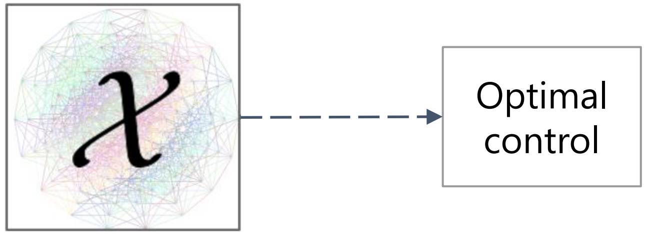
Control a high-dimensional Markov Decision Process

$$x_{t+1} = f_{\mathcal{X}}(x_t, u_t) + w$$

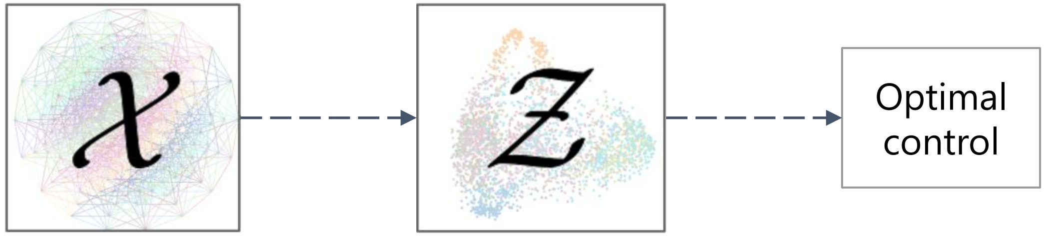


High-Dimensional Control is Challenging

Attempt optimal control in the observation space



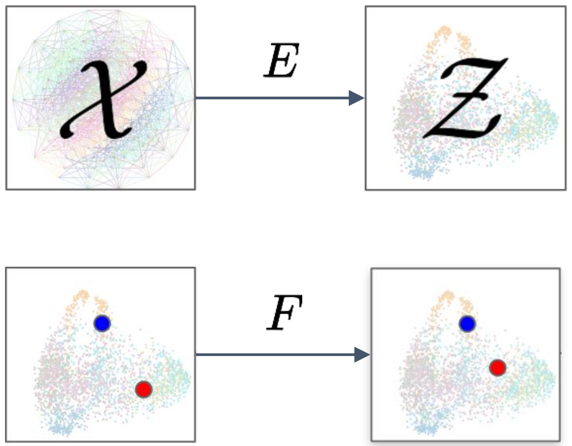
Embed then control



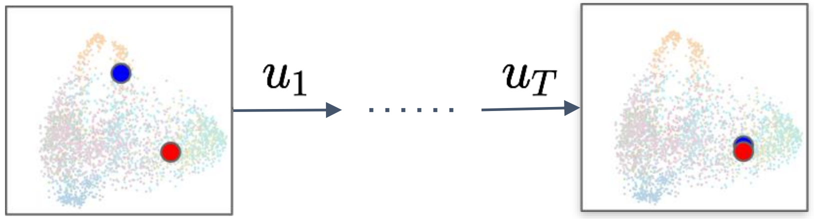
General approach

Learning Controllable Embedding (LCE): A two-step framework

Learn the embedding and latent dynamics

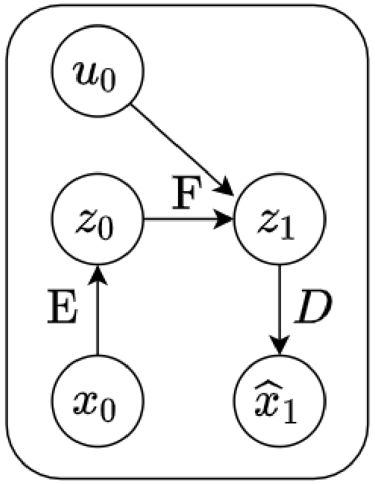


Control in the latent space

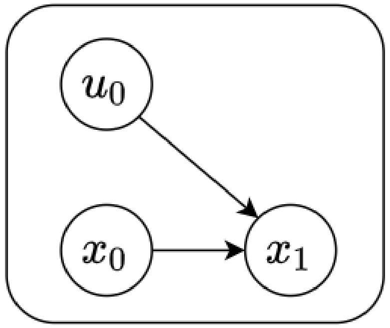


Existing LCE Models Use Explicit Prediction

Introduce an additional decoder D and learns to predict the next observation



$\approx$



Latent Variable Model

True environment dynamics

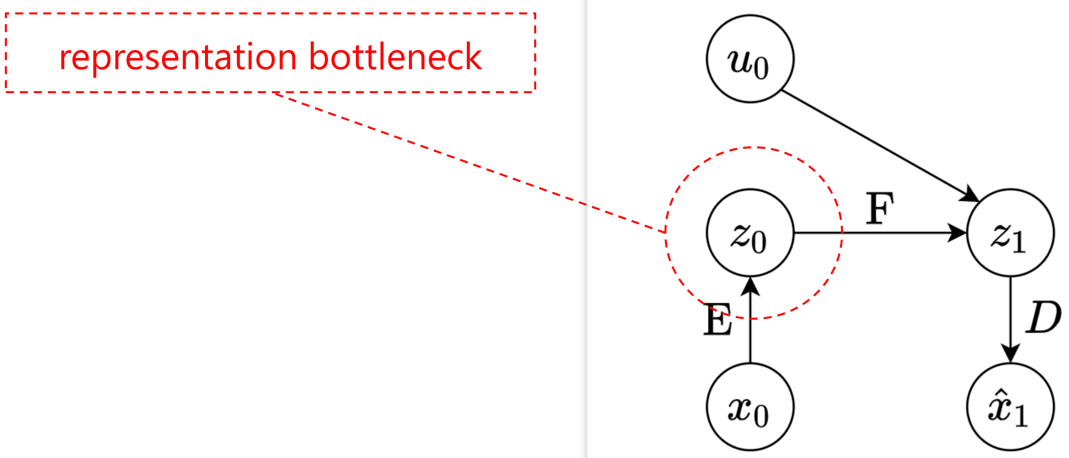
Contributions

Propose a decoder-free LCE-based model

Provide theoretical analysis for the learned embedding

Conduct extensive experiments

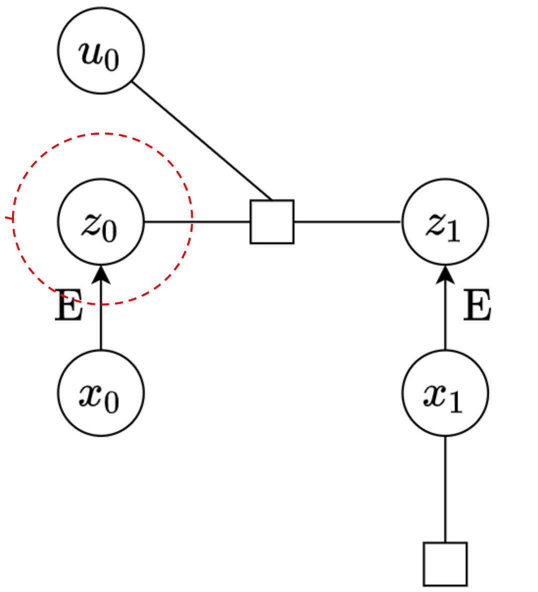
Learning an Embedding That is Good for Prediction



Decoder-based Directed Graph

Learning an Embedding That is Good for Prediction

representation bottleneck



Decoder-free Undirected Graph

Predictive suboptimality of a representation:

The *best possible* prediction loss achievable by such this model family for a fixed choice of representation

Learning an Embedding That is Good for Prediction

Lemma 1: Predictive suboptimality is upper bounded by the mutual information gap:

$$\underbrace{\boxed{I(x_{t+1}; x_t, u_t)}}_{\text{Information of observation data}} - \underbrace{\boxed{I(z_{t+1}; z_t, u_t)}}_{\text{Information captured in latent space}}$$

Predictive suboptimality can be minimized in a decoder-free and information-theoretic manner

Learning an Embedding via Predictive Coding

Predictive coding maximizes a lower bound of $I(z_{t+1}; z_t, u_t)$

critic

$$\ell_{\text{cpc}} = \mathbb{E}\frac{1}{K}\sum_i \ln \frac{\boxed{F}\left(E\left(x_{t+1}^{(i)}\right) | E\left(x_t^{(i)}\right), u_t^{(i)}\right)}{\frac{1}{K}\sum_j F\left(E\left(x_{t+1}^{(i)}\right) | E\left(x_t^{(j)}\right), u_t^{(j)}\right)}$$

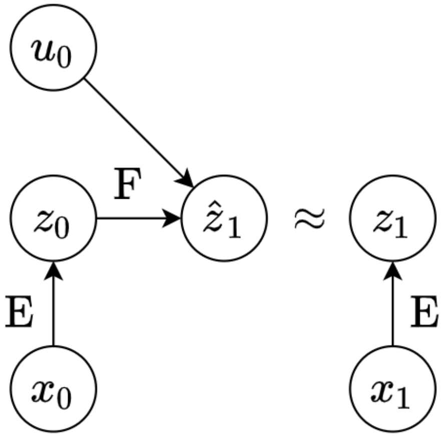
For a fixed choice of E, if F is the true latent dynamics, the CPC objective is maximized

Estimating the Latent Dynamics via Consistency

Choosing F to be the true latent dynamics maximizes the CPC objective

Choosing an F that maximizes the CPC objective does not imply F is the true latent dynamics model

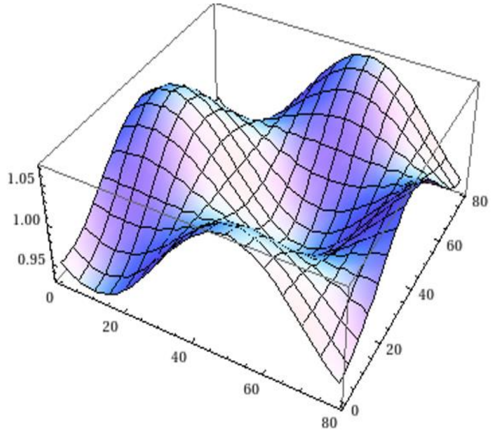
Explicitly encourage F to estimate the true latent dynamics via a *consistency objective*

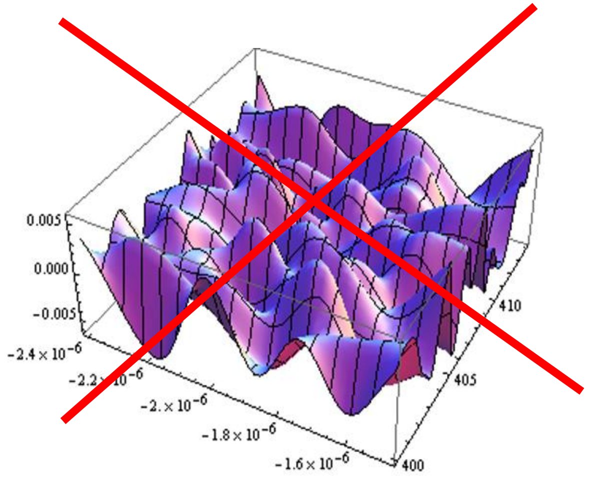


$$\ell_{\text{cons}} = \ln F(E(x_{t+1}) | E(x_t), u_t)$$

Enforcing Smooth Latent Dynamics via Curvature

We use locally-linear controllers (LLC), so a low-curvature dynamics is preferable





$$\ell_{\text{curv}} = \mathbb{E}_{\eta \sim \mathcal{N}(0,\delta I)} \Big[\| f_{\mathcal{Z}}(\bar{z},\bar{u}) - (\nabla_z f_{\mathcal{Z}}(\bar{z},\bar{u})\eta_z + \nabla_u f_{\mathcal{Z}}(\bar{z},\bar{u})\eta_u) - f_{\mathcal{Z}}(z,u) \|_2^2 \Big]$$

Predictive Coding, Consistency, Curvature (PC3)

$$\ell_{\text{cpc}} = \mathbb{E}\frac{1}{K}\sum_i \ln \frac{F\left(E\left(x_{t+1}^{(i)}\right)|E\left(x_t^{(i)}\right),u_t^{(i)}\right)}{\frac{1}{K}\sum_j F\left(E\left(x_{t+1}^{(i)}\right)|E\left(x_t^{(j)}\right),u_t^{(j)}\right)}$$

Maximize CPC

Learn good embedding E

$$\ell_{\text{cons}} = \ln F(E(x_{t+1})|E(x_t), u_t)$$

Maximize Consistency

Learn good dynamics F

$$\ell_{\text{curv}} = \mathbb{E}_{\eta\sim\mathcal{N}(0,\delta I)}\left[\|f_{\mathcal{Z}}(\bar{z},\bar{u}) - (\nabla_z f_{\mathcal{Z}}(\bar{z},\bar{u})\eta_z + \nabla_u f_{\mathcal{Z}}(\bar{z},\bar{u})\eta_u) - f_{\mathcal{Z}}(z,u)\|_2^2\right]$$

Minimize curvature

F **is suitable for LLC**

$$\boxed{\lambda_1\ell_{\text{cpc}} + \lambda_2\ell_{\text{cons}} + \lambda_3\ell_{\text{cur}}}$$

PCC vs PC3

Baseline

	PC3	PCC
Predictive Coding	✔	✘
Prediction	✘	✔
Consistency	✔	✔
Curvature	✔	✔
iLQR Controller	✔	✔

PC3 is contrastive analog of PCC

PCC vs PC3

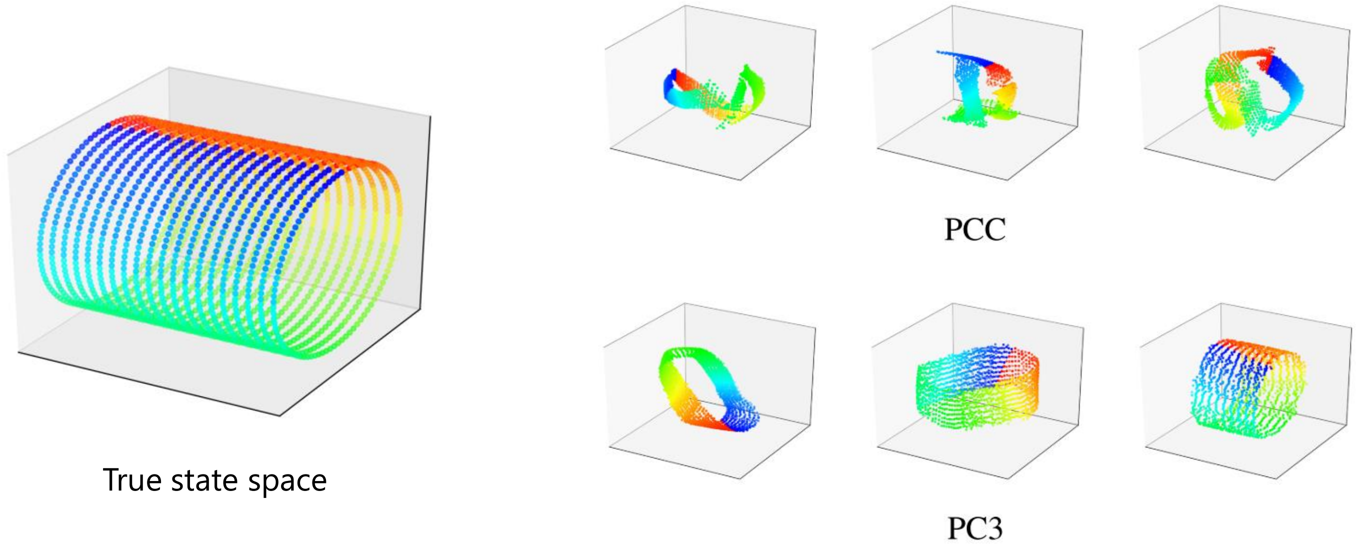
Control performance

Train 10 models, run 10 subtasks for each model (total 100 trials/method)

Task	PC3 (all)	PCC (all)	PC3 (top 1)	PCC (top 1)
Planar	**74.35 ± 0.76**	56.6 ± 3.15	**75.5 ± 0.32**	**75.5 ± 0.32**
Balance	**99.12 ± 0.66**	91.9 ± 1.72	**100 ± 0**	**100 ± 0**
Swing Up	**58.4 ± 3.53**	26.41 ± 2.64	**84 ± 0**	66.9 ± 3.8
Cartpole	**96.26 ± 0.95**	94.44 ± 1.34	**97.8 ± 1.4**	**97.8 ± 1.4**
3-link	**42.4 ± 3.23**	14.17 ± 2.2	**78 ± 1.04**	45.8 ± 6.4

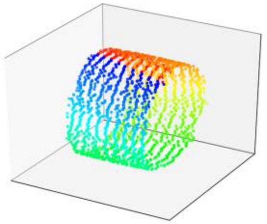
PCC vs PC3

Latent maps

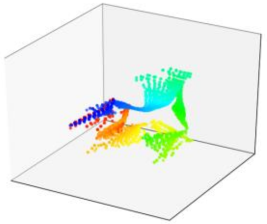


Ablation Analysis

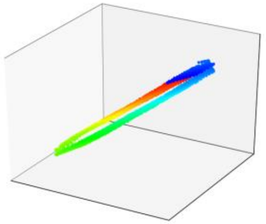
Setting	ℓ_{cpc}	ℓ_{cons}	ℓ_{cur}	Control
PC3	4.58	2.13	0.03	58.4 ± 3.53
w/o ℓ_{cons}	5.03	-4.87	0.0025	7.46 ± 1.32
w/o ℓ_{cur}	4.8	2.34	0.56	21.69 ± 2.73



PC3



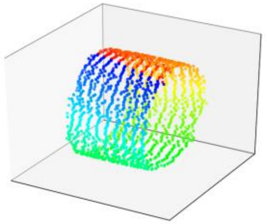
Without ℓ_{cons}



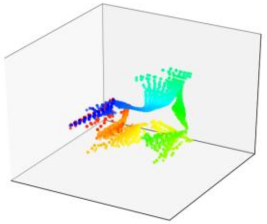
Without ℓ_{cur}

Ablation Analysis

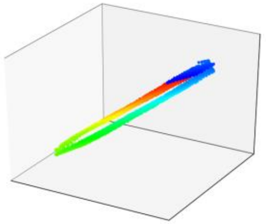
Setting	ℓ_{cpc}	ℓ_{cons}	ℓ_{cur}	Control
PC3	4.58	2.13	0.03	**58.4 ± 3.53**
w/o ℓ_{cons}	5.03	-4.87	0.0025	7.46 ± 1.32
w/o ℓ_{cur}	4.8	2.34	0.56	21.69 ± 2.73



PC3



Without ℓ_{cons}



Without ℓ_{cur}

Conclusion

- Predictive Coding-Consistency-Curvature (PC3) outperforms existing LCE-based models
- Promising information-theoretic extension of the LCE framework
- Future work
  - Scale to more complicated domains
  - Predictive coding in conjunction with other control algorithms

Thank you for listening!

Paper Link

www.vinai.io