Low-Rank Successor Representations Capture Human-Like Generalization

Eva Yi Xie, Nathaniel Daw*, Benjamin Eysenbach*
{evayixie, ndaw, eysenbach}@princeton.edu

Introduction

- 💡 Intelligent behavior hinges on predictive maps of future states, *aka* successor representations (SRs).
- 🤔 **Issue**: tabular SRs don't generalize to unseen states, *whereas* humans do!
- 💭 **Questions**:
  1. Do low-rank SRs support more efficient planning + broader generalization?
  2. If so, do they also capture human-like behavior bias?
- ✅ **Our results suggest Yes to both!**

Method

Function-approximated SR through Contrastive RL [1]
Given state set $\mathcal{S}$, learn low-d encoders $h_\theta, g_\theta: \mathcal{S} \to R^d$, $d \ll |\mathcal{S}|$. We define a *distance function*

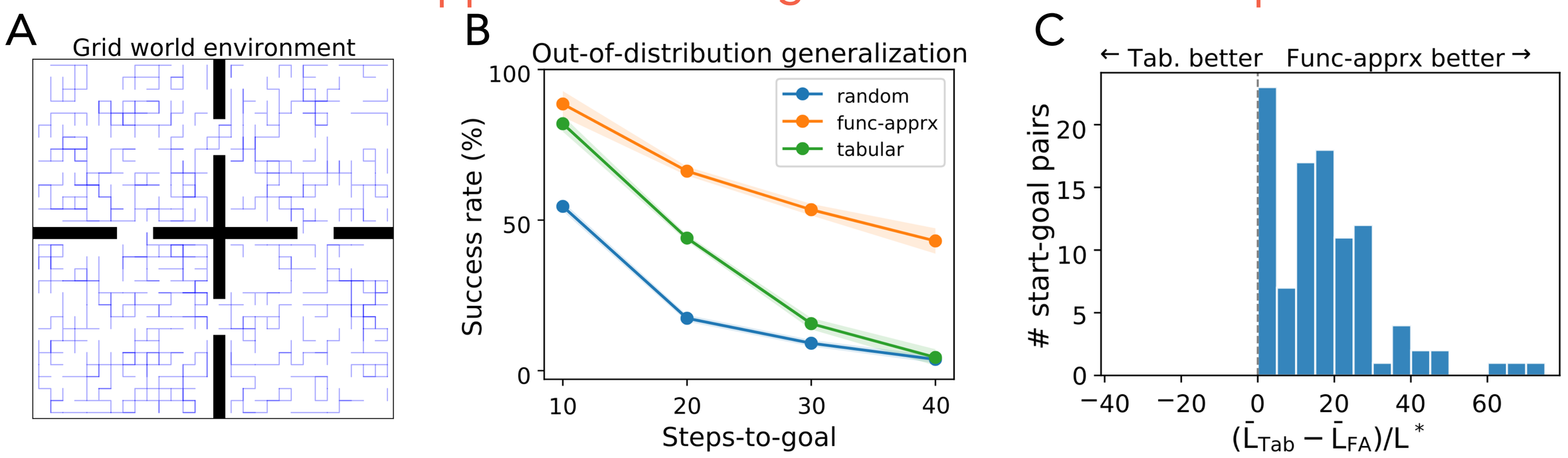
$$d_\theta(s, s') = h_\theta(\mathrm{s})^T g_\theta(s'), \text{and score } f_\theta(s, s') = -d_\theta(s, s').$$

The *forward InfoNCE loss* is

$$\mathrm{L}_{\text{forward}} = -E\left[\log \frac{\exp f_\theta(s_t, s_{t+\tau})}{\sum_{x \in \mathcal{N}_t} exp f_\theta(s_t, x)}\right],$$

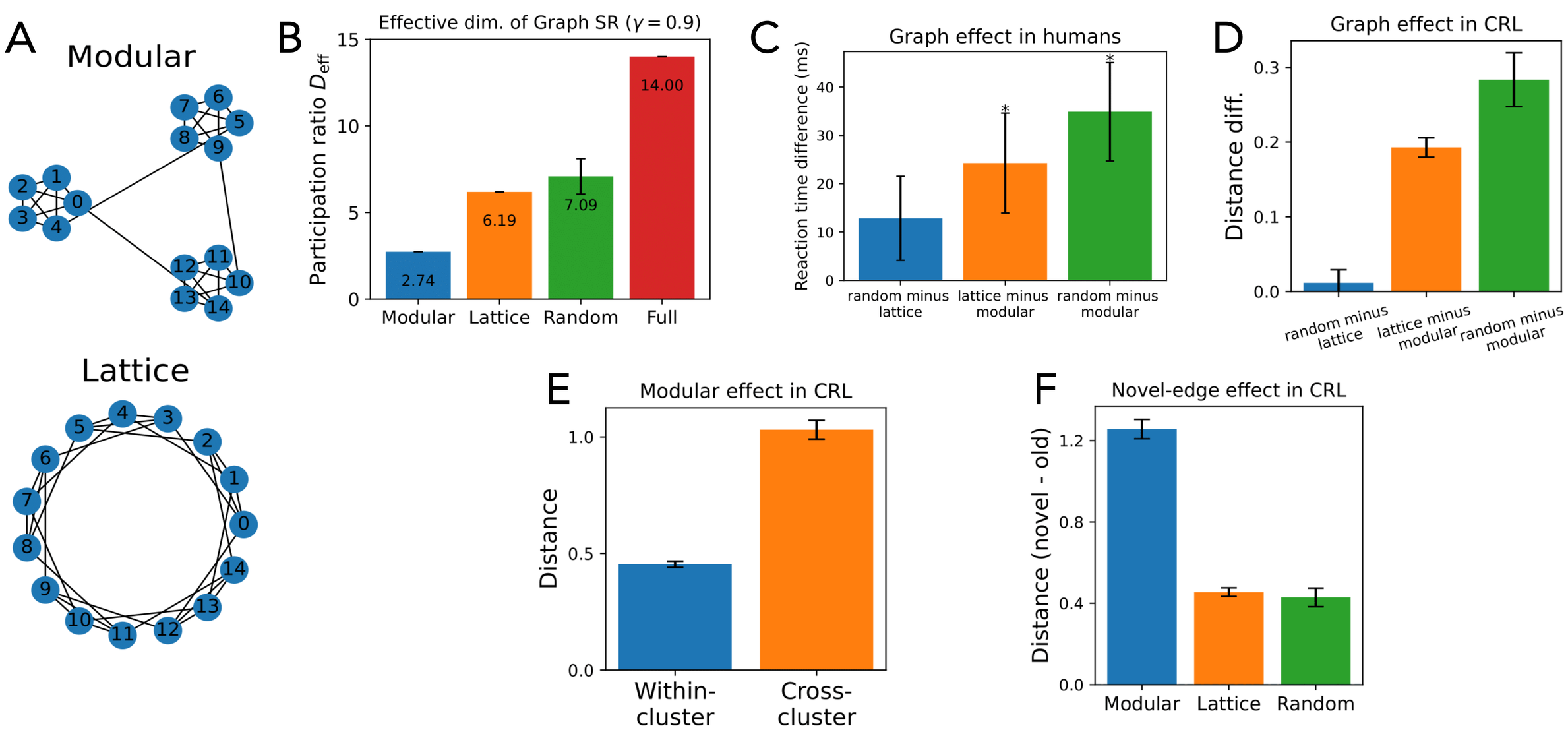
where $\mathcal{N}_t$ has one positive, $m-1$ negatives for anchor s_t.

Function-approximated SR generalizes & shortens paths



💡 **Fig 1.** We train on random length-5 trajectories in a grid world (1A) and found that function-approximated SR (FA-SR) generalizes to unseen states in evaluation. Among the start-goal pairs that tabular SRs could solve, FA-SA produces shorter paths.

Function-approximated SR is human-like in graphical tasks



💡 **Fig 2.** In high-d graphical tasks [2], subjects are shown random walks from an unknown graph type (2A, B). Humans take longer to react to random ≈ lattice > modular graphs (2C, adapted from [2]), captured by FA-SR (2D). FA-SR also captures other human behaviors (2E, F).

Conclusion + Next Steps

- ✅ FA-SR supports generalization to novel states.
- ✅ FA-SR plans shorter trajectories than full-rank SR.
- ✅ FA-SR shapes human-like behavior bias qualitatively.
- 💡 Directly fit human reaction time data [2] with FA-SR.
- 💡 Extend grid world to richer, more naturalistic environments.
- 💡 Involve neural data, e.g., hippocampal-entorhinal representations.

References:
1. Benjamin Eysenbach, Tianjun Zhang, Sergey Levine, and Russ R. Salakhutdinov. Contrastive learning as goal-conditioned reinforcement learning. *Advances in Neural Information Processing Systems*, 35:35603–35620, 2022.
2. Ari E. Kahn, Elisabeth A. Karuza, Jean M. Vettel, and Danielle S. Bassett. Network constraints on learnability of probabilistic motor sequences. *Nature human behaviour*, 2(12):936–947, 2018.