Contrastive Difference Predictive Coding

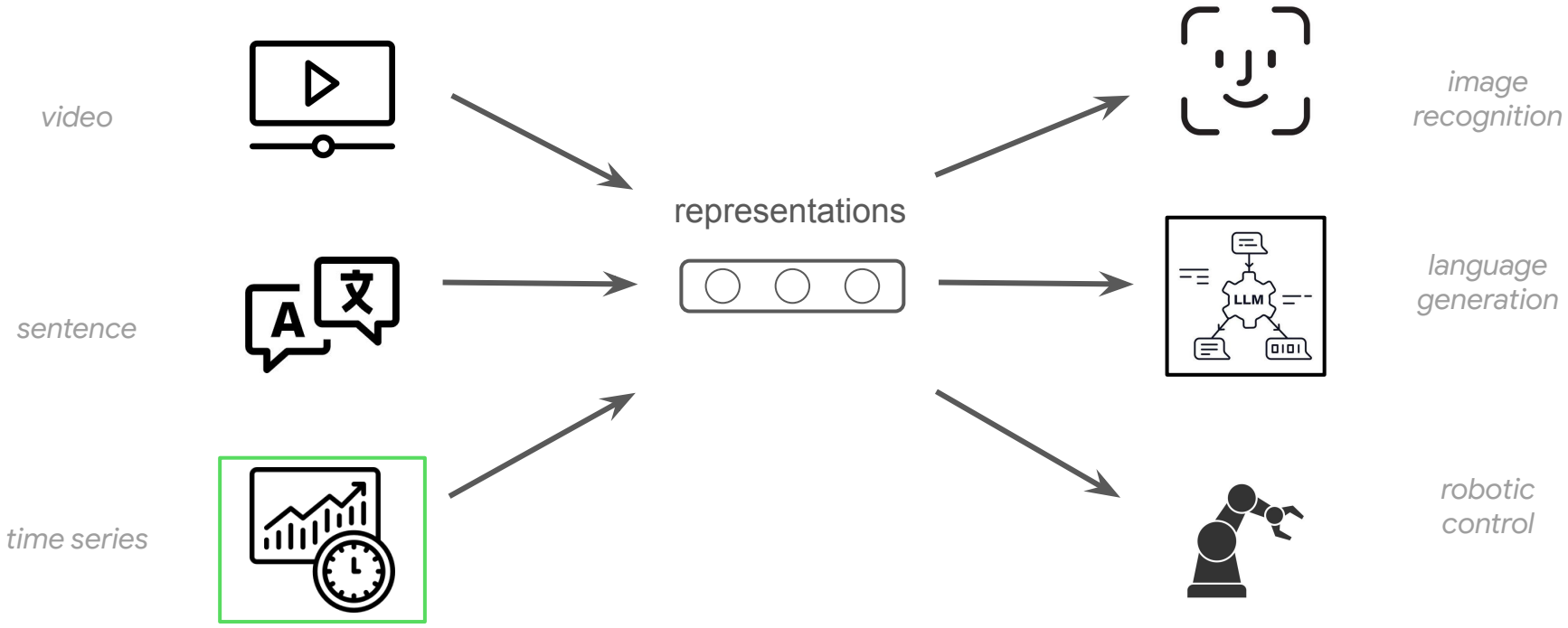
Chongyi Zheng, Ruslan Salakhutdinov, Benjamin Eysenbach

chongyiz@andrew.cmu.edu

Carnegie Mellon University

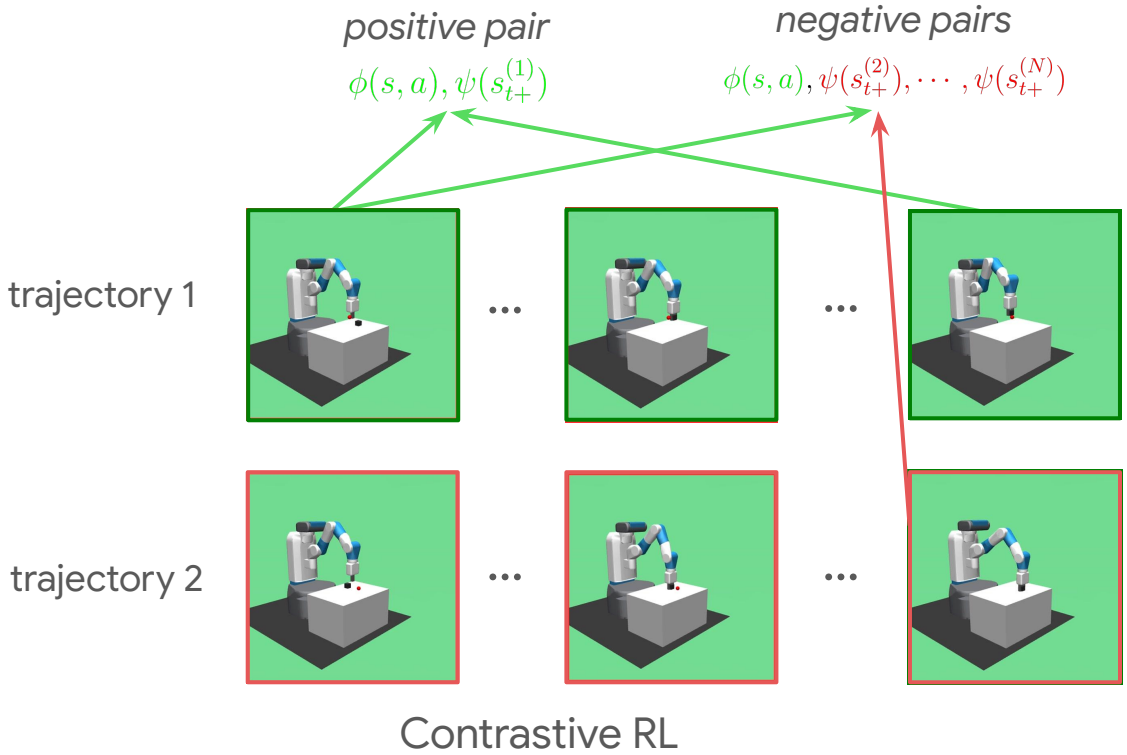
PRINCETON UNIVERSITY

Representation learning promises to solve different tasks.



[1] He et al. Masked autoencoders are scalable vision learners. 2022
[2] Gao et al. SimCSE: Simple Contrastive Learning of Sentence Embeddings. 2021
[3] Oord et al. Representation learning with contrastive predictive coding. 2018

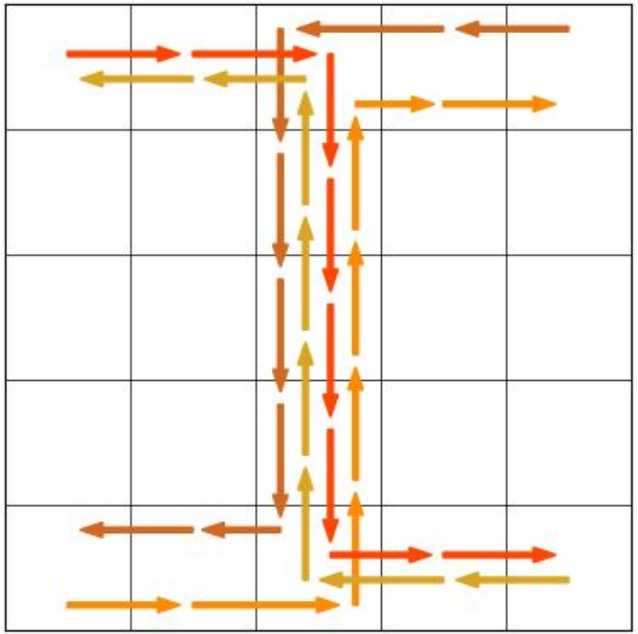
Contrastive RL for goal-reaching problems



$$p^{\pi}(s_{t+} \mid s, a)$$

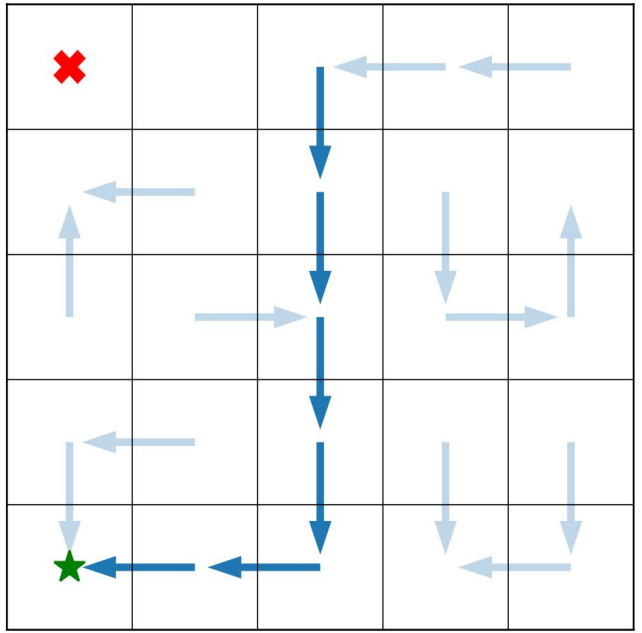
[1] Eysenbach et al. Contrastive Learning as Goal-Conditioned Reinforcement Learning. 2022
[2] Hénaff. Data-efficient image recognition with contrastive predictive coding. 2020

MC InfoNCE fails to do combinatorial generalization. (didactic example)



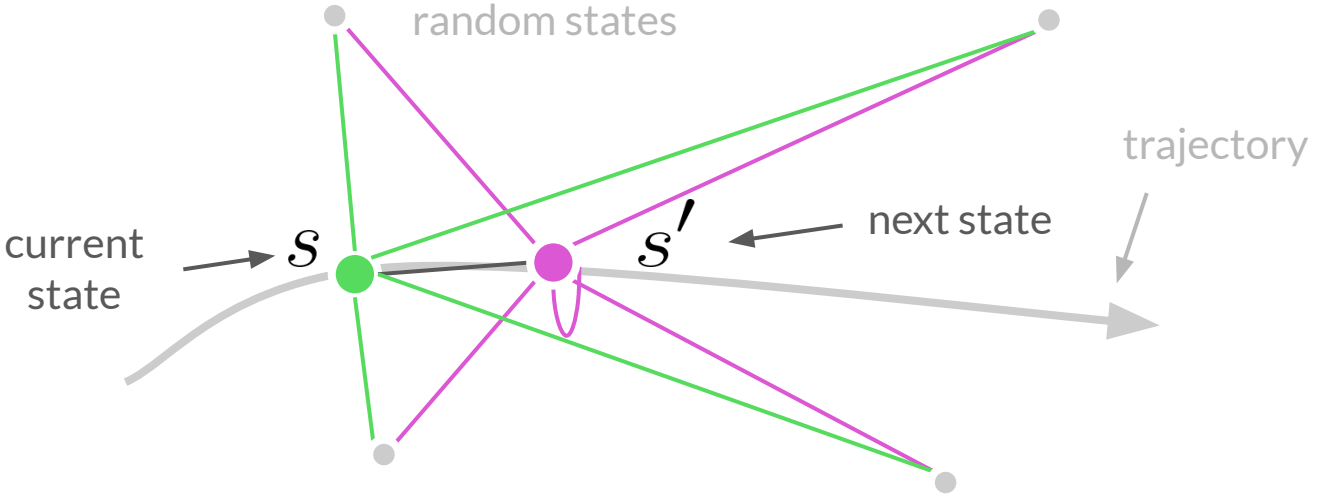
start ✖

MC InfoNCE



goal ★

Idea of temporal difference (TD) InfoNCE



$$p^{\pi}(s_{t+} = \cdot \mid s, a) = (1-\gamma)\,\delta(s' = \cdot) + \gamma\, p^{\pi}(s_{t+} = \cdot \mid s', a')$$

$$a' \sim \pi(a' \mid s')$$

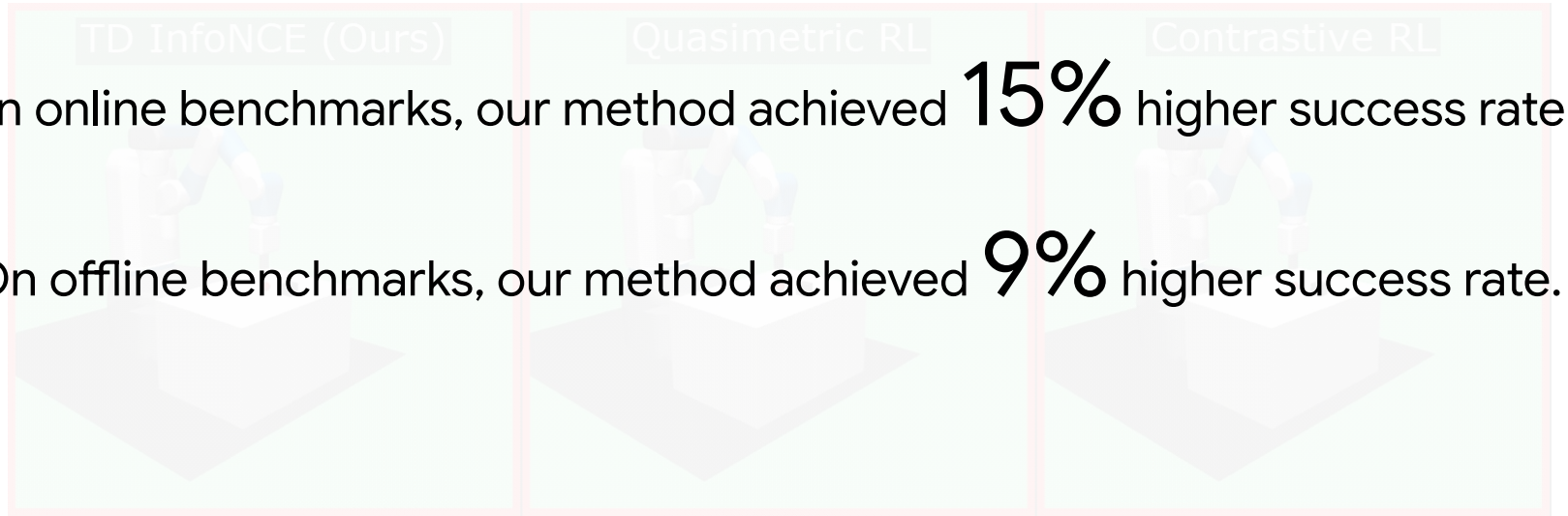
TD InfoNCE enables combinatorial generalization.

In tabular settings, our method increases sample efficiency by up to

1500x

Solving goal-conditioned manipulation tasks

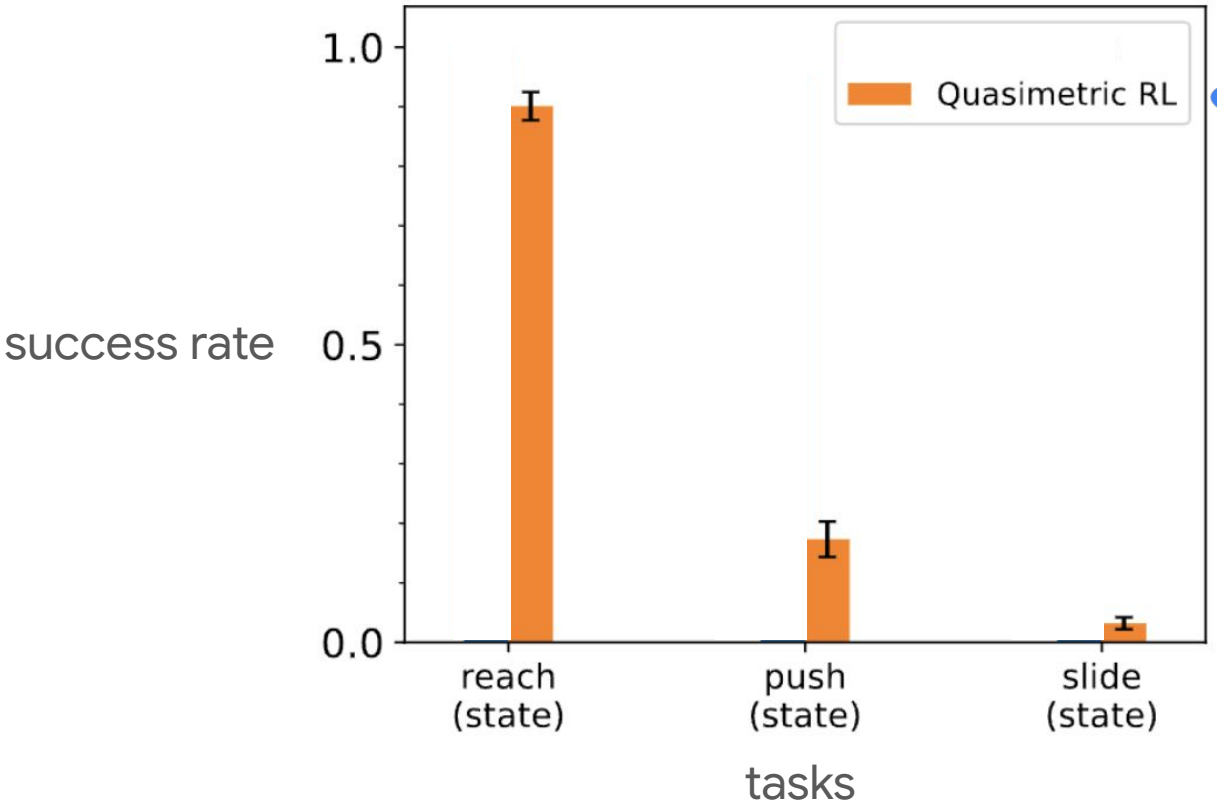
pick and place



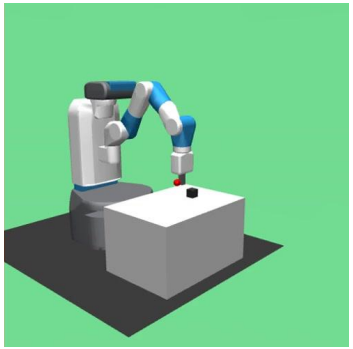
On online benchmarks, our method achieved 15% higher success rate.

On offline benchmarks, our method achieved 9% higher success rate.

TD InfoNCE is especially efficient in stochastic settings.



Prior method implicitly assumes that dynamics are deterministic.



Connections and future directions.

tldr: TD learning for temporal contrastive learning

- ➔ Big boost in sample efficiency
- ➔ Off-policy (i.e., counterfactual) reasoning
- ➔ Connections with many other areas:
  - ◆ Provably related to Q-values for arbitrary tasks
  - ◆ Mutual information
  - ◆ Successor representations
- ➔ Many opportunities for future work!

Video, code, and paper!

https://chongyi-zheng.github.io/td_infonce