

Autonomous Learning with Goals

Goal conditioning and goal selection without external rewards in RL

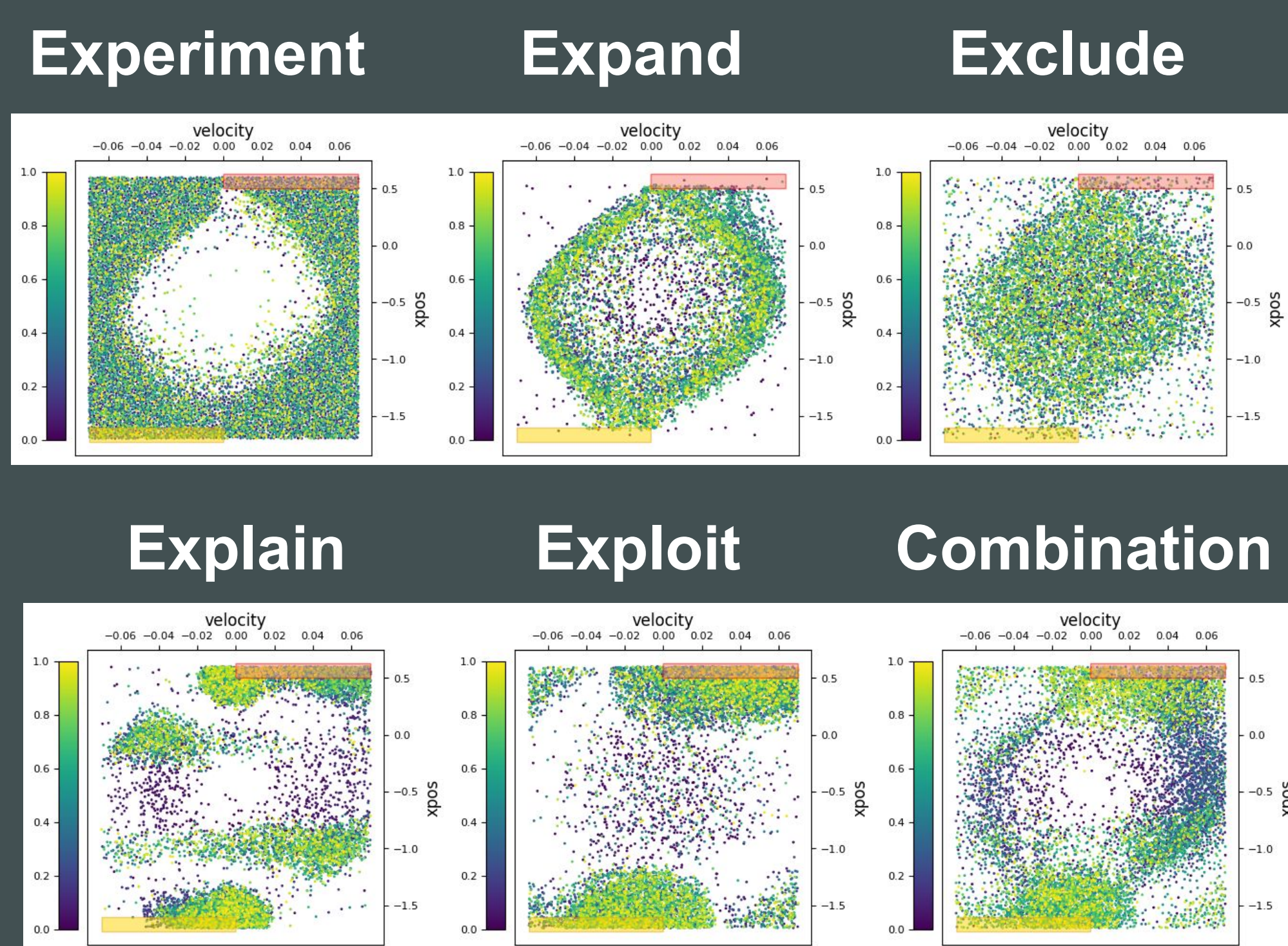
Hampus Åström, Lund University, Computer Science

Goal conditioning can turn any environment into an open learning environment, without external reward.
Skills represented as goals, can be acquired autonomously and in an explainable format.

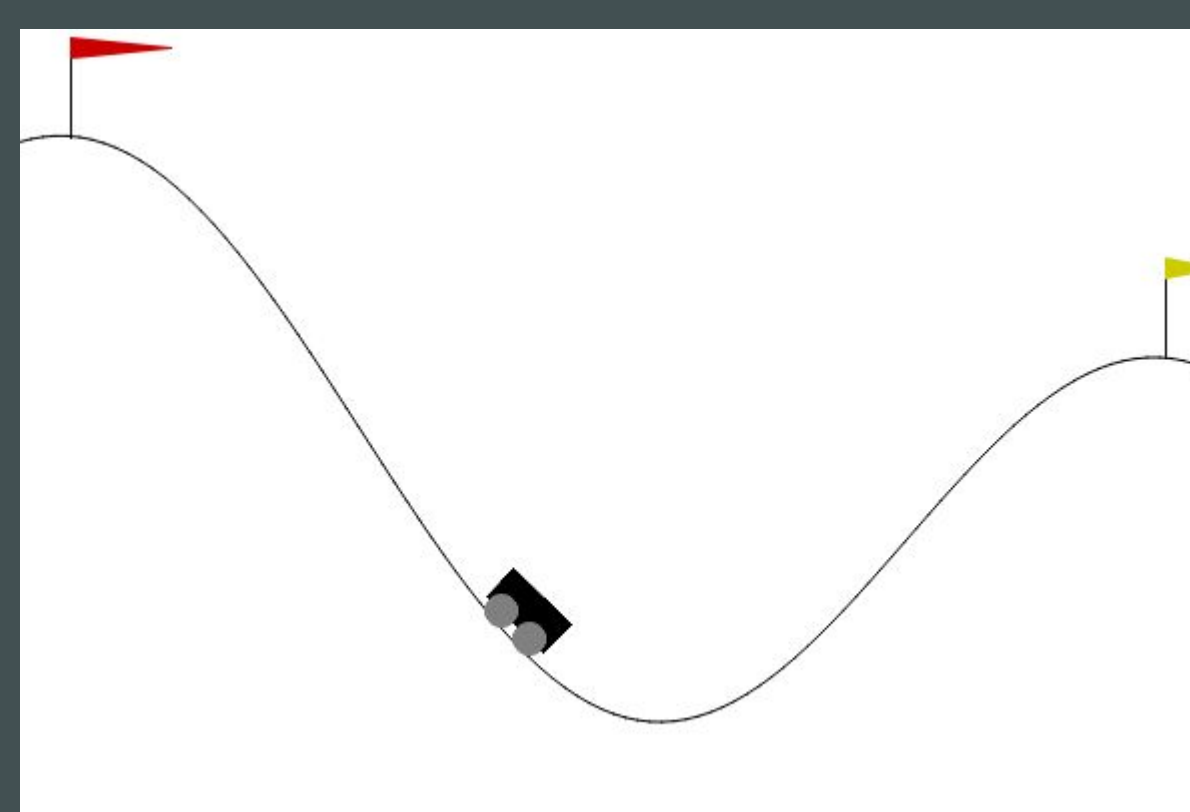
How are goals best selected for exploration and stability?

Past: Goal Exploration,
a study in distance-based goal selection

Goal spread with different goal selection methods:



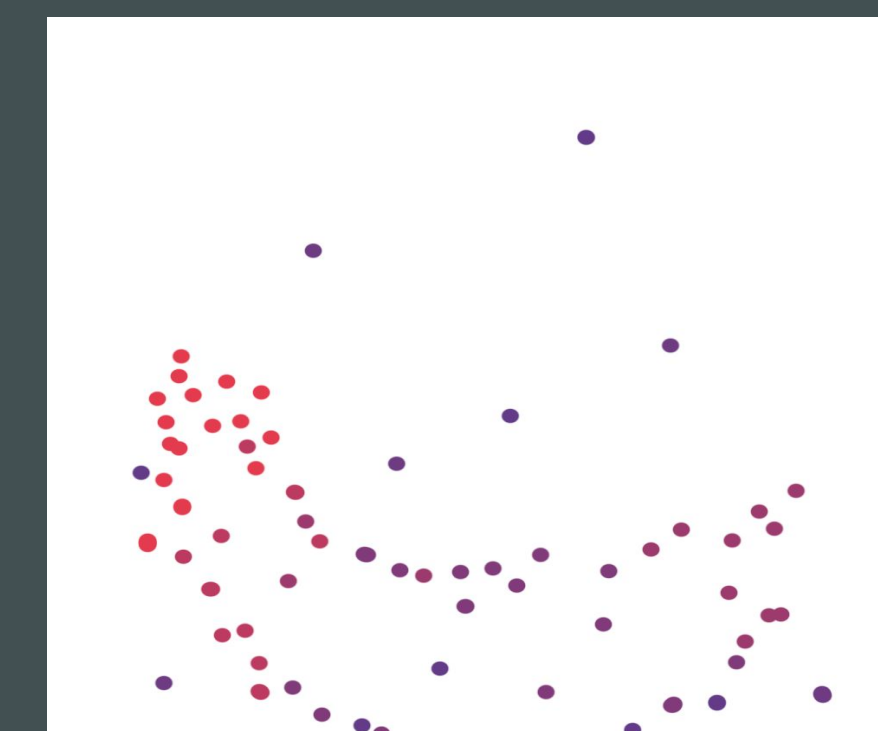
Pathological Mountain
Car Environment



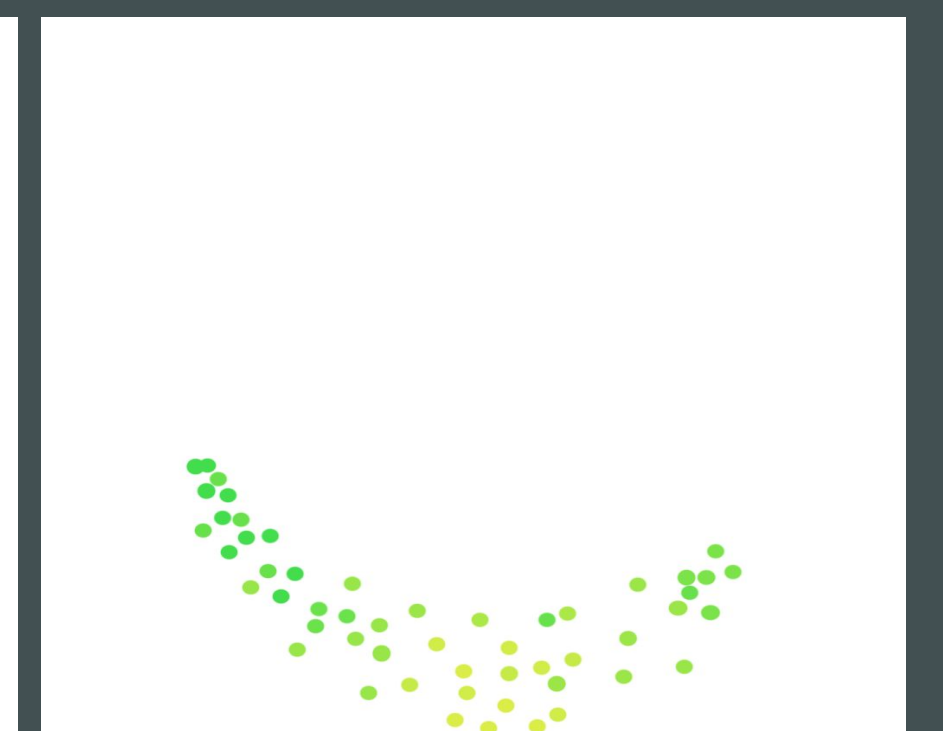
Environment inspired by:
S. Chakraborty et al., "Dealing with Sparse Rewards in
Continuous Control Robotics via Heavy-Tailed Policy
Optimization," 2023, ICRA

Future: Exploration and Stability,
with GAN and Mapping?

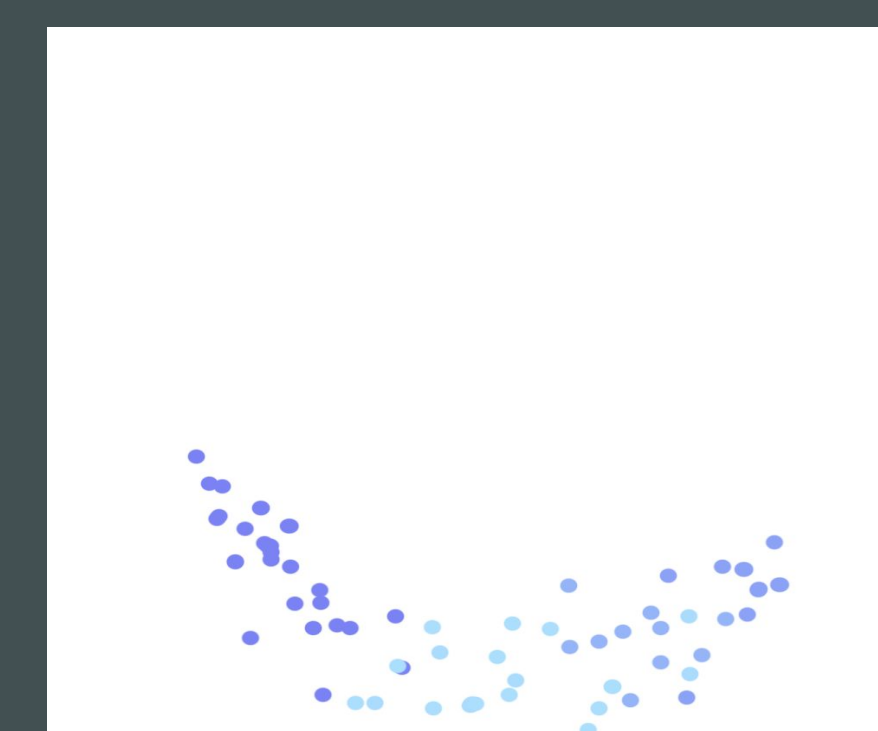
Failed goals



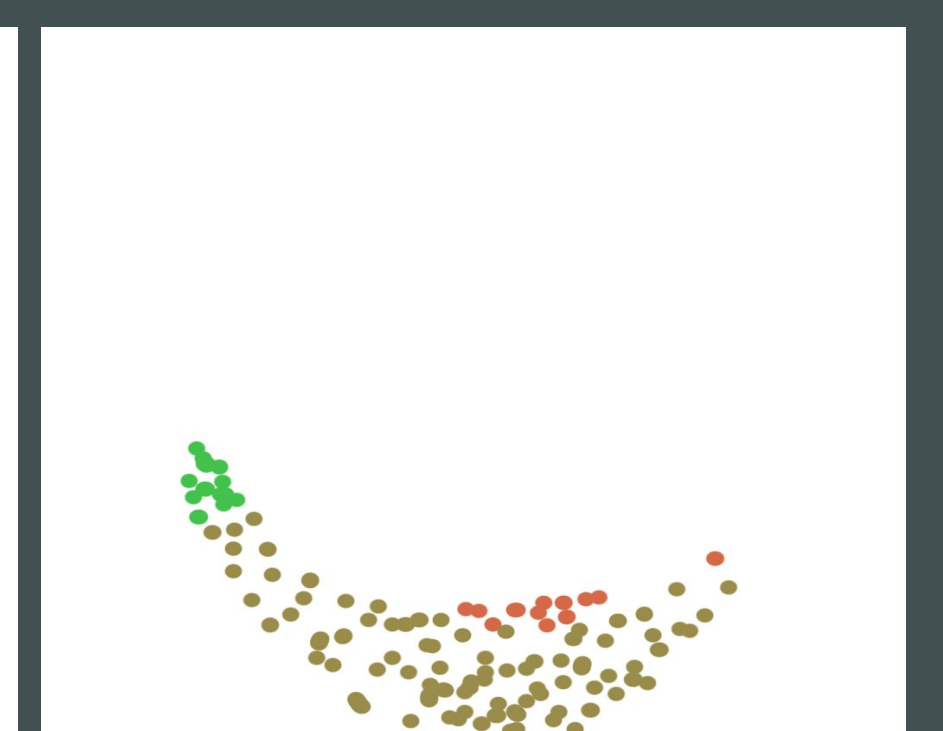
Successful goals



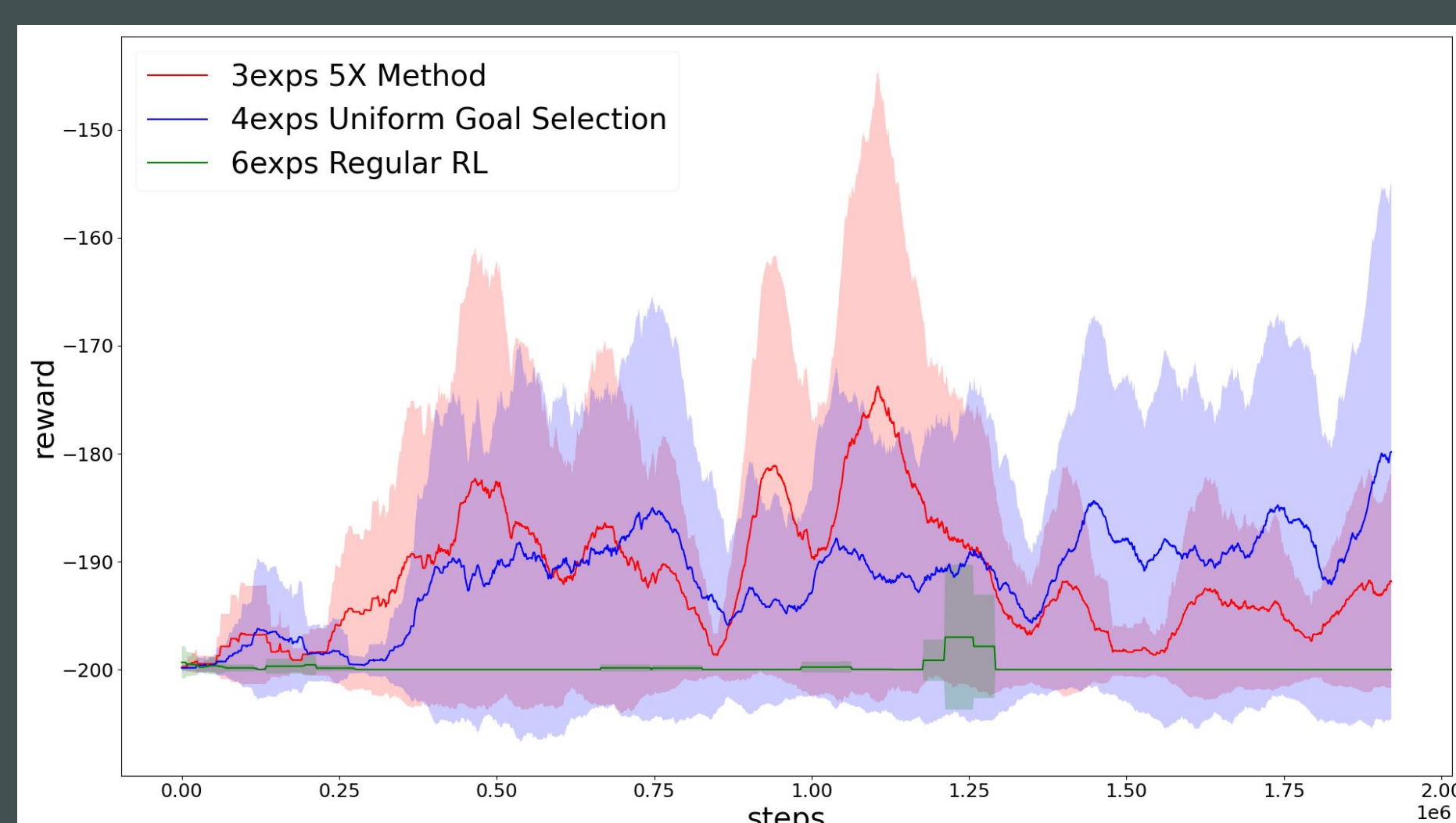
Intrinsic rewards



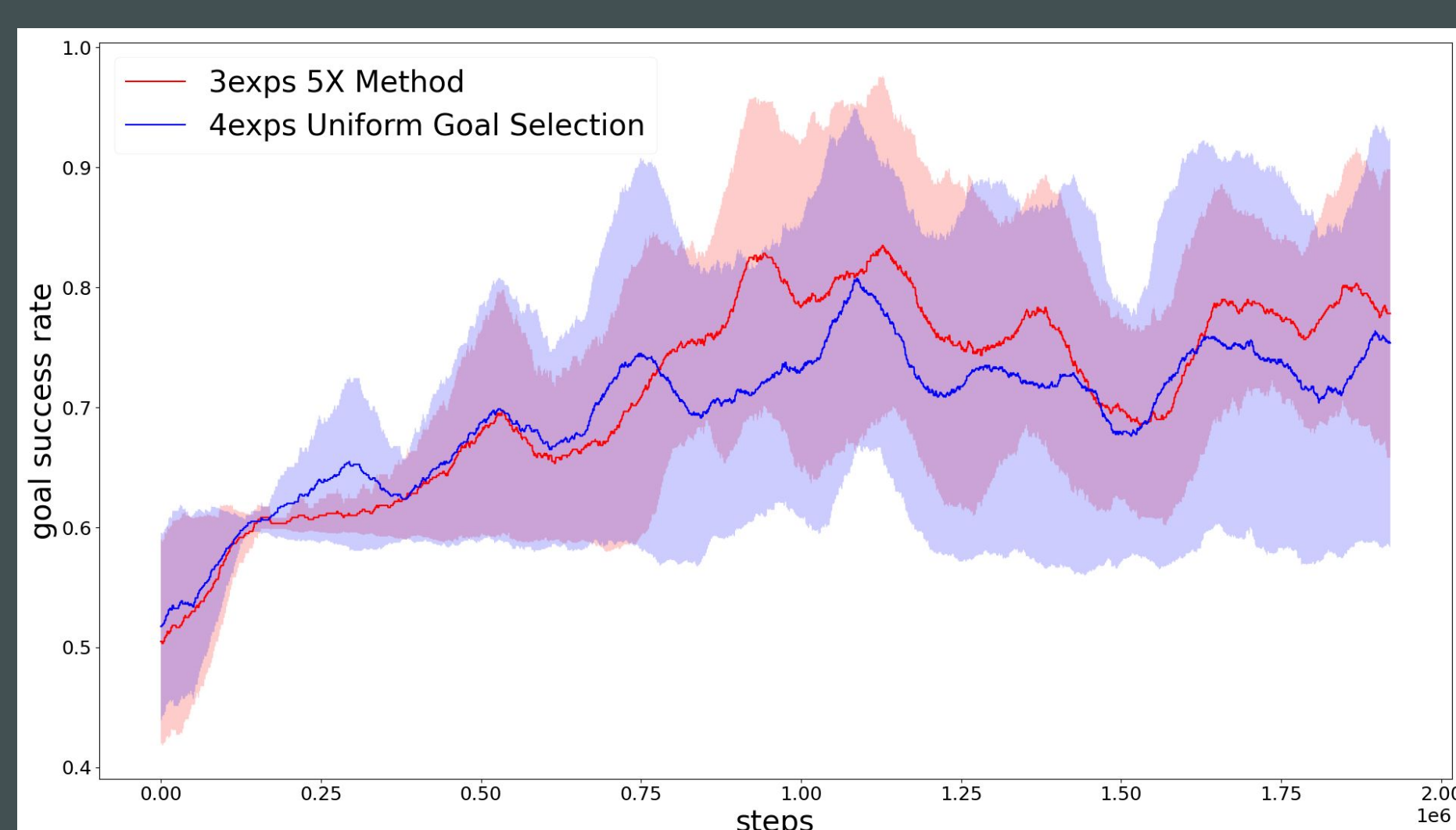
Extrinsic rewards



External reward in
evaluation

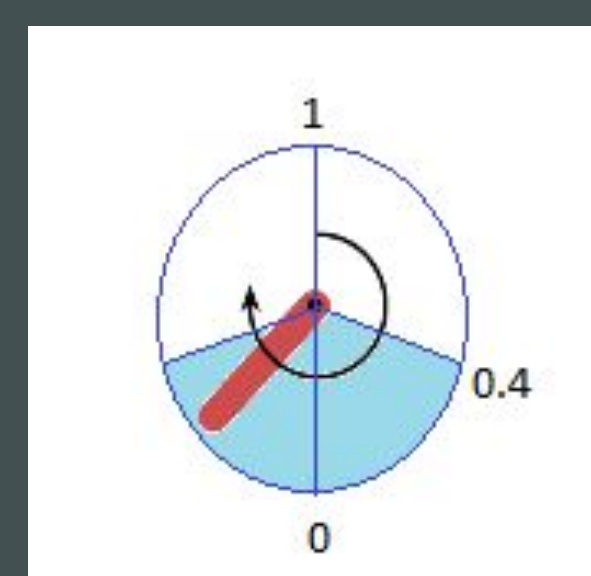


Goal reaching for
set of test goals

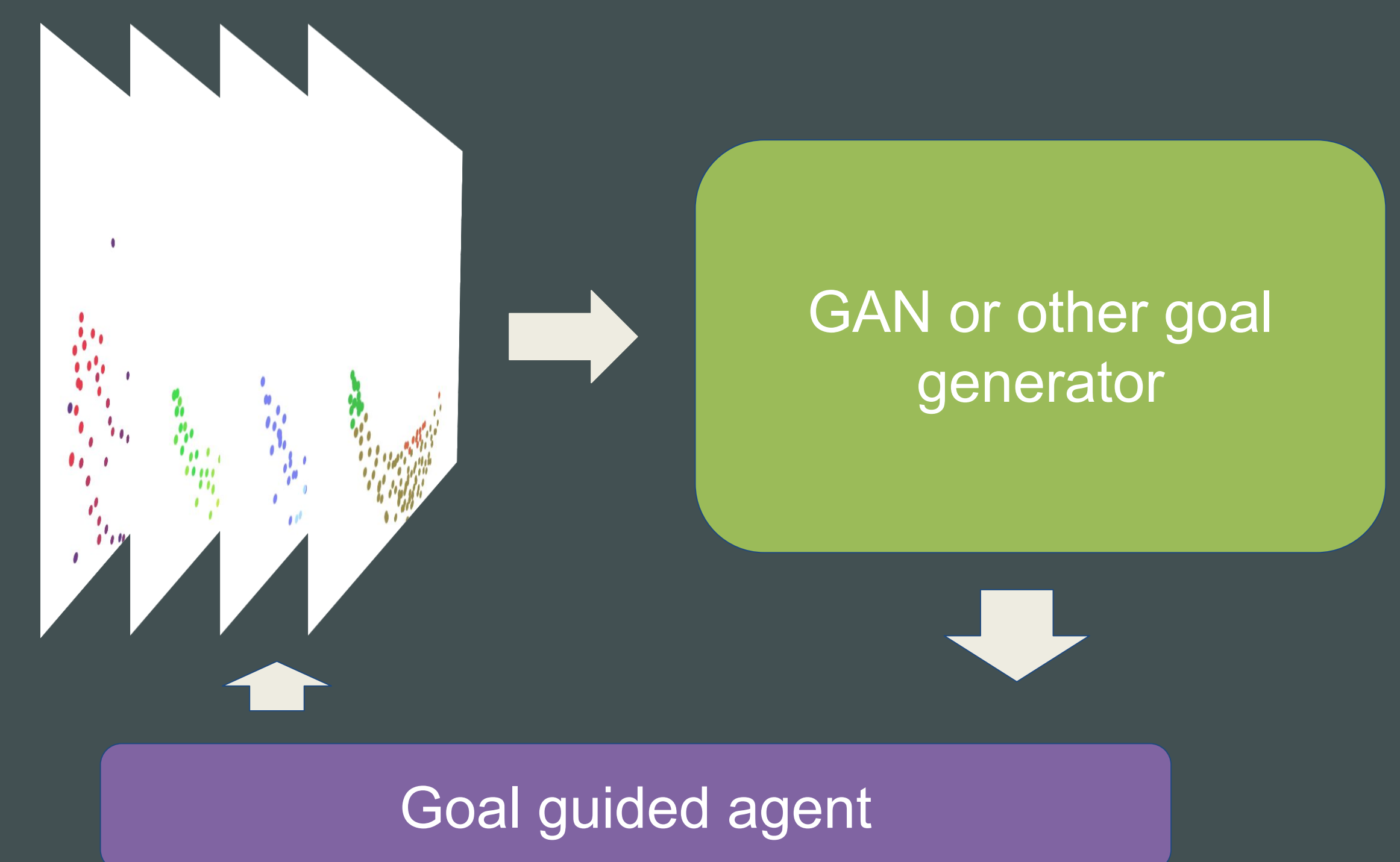


Goal Selection strategies are hard to evaluate
due to instability (high standard deviations)

Sparse Pendulum
Environment



Sketch of goal and observation mappings,
most decayed over time (not real data).



A GAN (or other goal generator) selects goal,
agent pursues goal and collects data,
data is used in goal selection

If you want to use my wrapper for turning any environment into a goal environment, contact me! hampus.astrom@cs.lth.se

Autotelic Learning with Hindsight

Intrinsic motivation, e.g. curiosity [1], guides exploration without external reward or goals.

Goals and hindsight learning [2] can guarantee a reward signal from each episode.

Autotelic learning [3] - the agent selects its goals itself, to maximize its own learning metric.

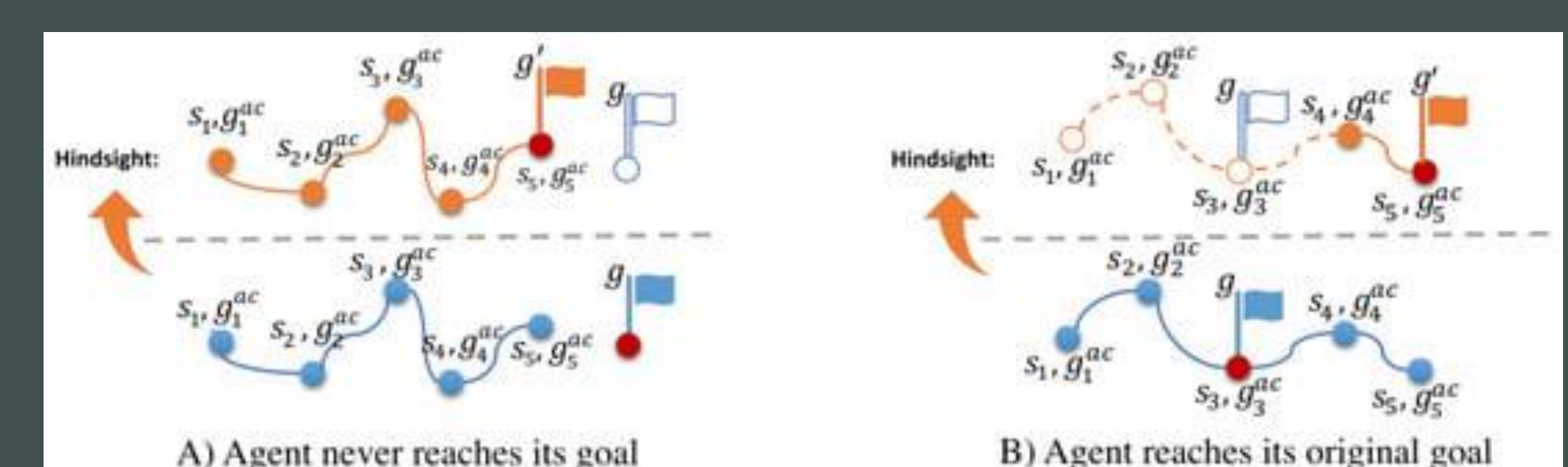


Image taken from T. Dai, H. Liu, A. Anthony Bharath "Episodic Self-Imitation Learning with Hindsight." in Electronics. 2020

References

- [1] D. Pathak, P. Agrawal, A. A. Efros, and T. Darrell, "Curiosity-Driven Exploration by Self-Supervised Prediction," in 2017 IEEE (CVPRW), Honolulu, HI, USA
- [2] M. Andrychowicz et al., "Hindsight Experience Replay." arXiv, Feb. 23, 2018.
- [3] C. Colas, T. Karch, O. Sigaud, and P.-Y. Oudeyer, "Autotelic Agents with Intrinsically Motivated Goal-Conditioned Reinforcement Learning: A Short Survey," Journal of Artificial Intelligence Research, Jul. 2022