Explore and Control via Adversarial Surprise

Paper: arxiv.org/abs/2107.07394

Effective unsupervised reinforcement learning requires a balance between seeking novelty and familiarity. How can we build an algorithm that strikes this balance?

Previous unsupervised RL approaches generally fall into one of two categories: those that seek novelty and those that seek familiarity. However, optimizing only one or the other is not effective. We propose to reconcile these seemingly opposing objectives as an adversarial game.

Adversarial Surprise produces complex and meaningful behaviors by formulating an adversarial game between two policies acting on the environment. The force of our approach is the information-theoretic formulation of the game which makes it general and theoretically sound.

We apply this method to the control of a single embodied RL agent, where the two policies compete over the amount of surprise the agent experiences.
From this internal competition emerge complex behaviors without any external supervision.
Adversarial Surprise produces meaningful behaviors without external reward in two popular benchmarks: Atari and VizDoom.

We show that we get emergent complexity by exhibiting clear phase transitions.

We procedurally generate custom MiniGrid environments to test the coverage ability of unsupervised RL methods in stochastic partially-observable environments. These environments contain noisy rooms with controllable noise and dark rooms.

RND gets distracted by the stochastic elements of the noisy rooms ("noisy TV" problem) and SMiRL stays in the nearest niche it finds ("dark room" problem). The increasingly complex exploration and control strategies produced by Adversarial Surprise solve both problems.

We show that under certain conditions, Adversarial Surprise provably covers the latent state space of the environment.