Paper summary: Never give up: learning directed exploration strategies

March 20, 2022

1 Idea in few sentances

2 Explanation of the central concept

3 Methodology

4 Initial rambly notes

4.1 Abstract

A RL designed to solve hard exploration problems is proposed. The idea is to learn a range of directed exploratory policies. An episodic memory-based intrinsic reward is constructed using k-nearest neighbors over the agent's recent experience. That reward is used to train the directed exploratory policies, thereby ecouraging the agent to repeatedly revisit all states in its environment. A self-supervised inverse dynamics model is used to train the embedding of the nearest neighbor lookup so that the novelty is biased toward what the agent can control. Universal value function approximators (UVFA) are used to simultaneously learn the different directed exploratory policies within the same neural network.

Apparently i need to go over transfer and inverse reinforcement learning to read this.................. Well back to the berkley lectures then.

4.2 Introduction

4.3 Method

4.4 Other stuff