CSC2621 Topics in Robotics
Reinforcement Learning in Robotics

Week 6: Distributional RL

Animesh Garg

A distributional code for value in dopamine based reinforcement learning

Dabney et al. Nature 2020

Topic: Distributional RL
Presenter: Animesh Garg

Will Dabney[1,5]*, Zeb Kurth-Nelson[1,2,5], Naoshige Uchida[3], Clara Kwon Starkweather[3], Demis Hassabis[1], Rémi Munos[1] & Matthew Botvinick[1,4,5]

Since its introduction, the reward prediction error theory of dopamine has explained a wealth of empirical phenomena, providing a unifying framework for understanding the representation of reward and value in the brain[1–3]. According to the now canonical theory, reward predictions are represented as a single scalar quantity, which supports learning about the expectation, or mean, of stochastic outcomes. Here we propose an account of dopamine-based reinforcement learning inspired by recent artificial intelligence research on distributional reinforcement learning[4–6]. We hypothesized that the brain represents possible future rewards not as a single mean, but instead as a probability distribution, effectively representing multiple future outcomes simultaneously and in parallel. This idea implies a set of empirical predictions, which we tested using single-unit recordings from mouse ventral tegmental area. Our findings provide strong evidence for a neural realization of distributional reinforcement learning.

Motivation

Rewards are not Scalar

"According to the now canonical theory, reward predictions are represented as a single scalar quantity, which supports learning about the expectation, or mean, of stochastic outcomes..."

Multiplicity of Future Outcomes

"We hypothesized that the brain represents possible future rewards not as a single mean, but instead as a probability distribution, effectively representing multiple future outcomes simultaneously and in parallel...."

Motivation
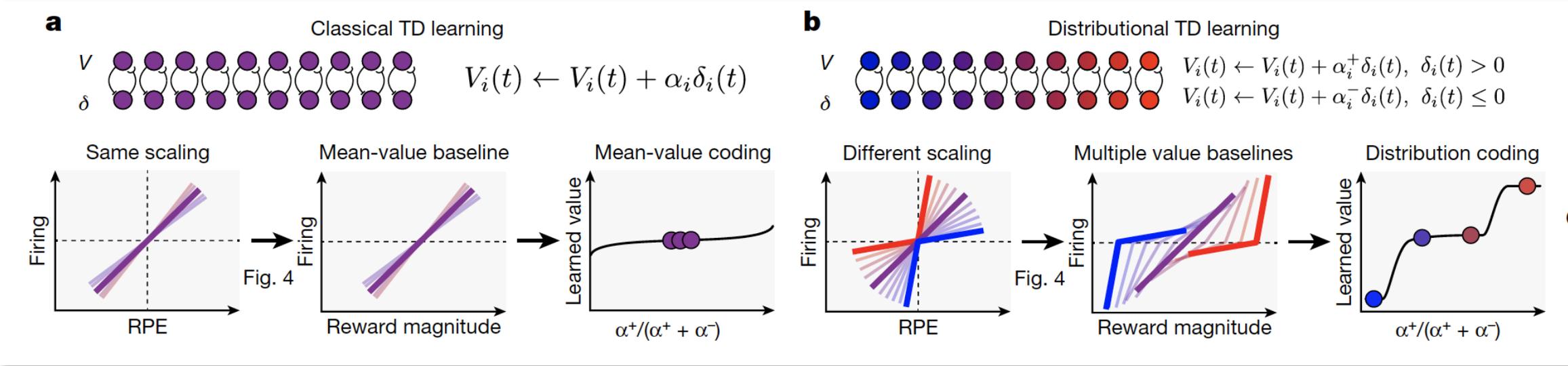
Distributional value coding arises from a diversity of relative scaling of positive and negative prediction errors

Agenda

Q-Value (Continued)

- QT-Opt: Scalable Deep Reinforcement Learning for Vision-Based Robotic Manipulation
  Make distributed Q-learning (+variants) work for real applications

Distributional RL

- A Comparative Analysis of Expected and Distributional Reinforcement Learning
  What is Distributional RL? Does any of this work at all? If yes, then when?

- Statistics and Samples in Distributional Reinforcement Learning
  How and when Categorical and Quantile regression is insufficient. Hence Expectiles!