

Sparse Coding of Learned State Representations in Reinforcement Learning

Submission ID 3000324
Submission Type Poster
Topic Cognitive Science
Status Submitted
Submitter David Noelle
Affiliation University of California, Merced

SUBMISSION DETAILS

Presentation Type Either Poster or Oral Presentation

Presentation Abstract Summary Temporal Difference (TD) Learning is a leading account of the role of the dopamine system in reinforcement learning. TD Learning has been shown to fail to learn some fairly simple control tasks, however, challenging this explanation of reward-based learning. We conjecture that such failures do not arise in the brain because of the ubiquitous presence of lateral inhibition in the cortex, producing sparse distributed internal representations that support the learning of expected future reward. We provide support for this position by demonstrating the benefits of learned sparse representations for two problematic control tasks: mountain car and acrobat.

Co-author Information

* Presenting Author

| First Name | Last Name | Affiliation | E-mail |
|------------|-----------|----------------------------------|----------------------------|
| David * | Noelle * | University of California, Merced | dnoelle@ucmerced.edu |
| Jacob | Rafati | University of California, Merced | jrafatiheravi@ucmerced.edu |

Keywords

| Keywords |
|------------------------|
| reinforcement learning |
| sparse coding |
| dopamine |
| lateral inhibition |

value function approximation