Human Adaptation of Learning Strategies Resembles Policy Gradients

Hua-Dong Xiong
Georgia Institute of Technology, Atlanta, Georgia, United States

JIAN LI
University of California San Diego, La Jolla, California, United States

Marcelo Mattar
New York University, New York, New York, United States

Robert Wilson
Georgia Tech, Atlanta, Georgia, United States

Abstract

A hallmark of human intelligence is not only the capacity to learn from the environment, but also the ability to adapt the learning process itself in response to changing demands. This meta-learning ability, known as "learning to learn," has been extensively studied in cognitive science and artificial intelligence for decades. While task-optimized recurrent neural networks have offered qualitative accounts of biological learning-to-learn, they fall short in capturing the individual and temporal variability inherent in human decision making. To investigate how humans adjust their learning strategies over time, we introduce a neural network model that estimates dynamic changes in subjects' reinforcement learning (RL) parameters. Across four bandit tasks, we find that RL parameters change over time, indicating that humans continuously adapt their decision-making strategies at both the trial and block levels. These parameter updates are associated with greater rewards and align with policy gradients near current RL parameters, suggesting that humans refine their learning strategies based on task feedback. Taken together, our work provides a novel framework for understanding the adaptive mechanisms of biological meta-learning, with broad applicability across tasks, populations, and cognitive models.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*