Regression, encoding, control: an integrated approach to shared representations with distributed coding

Gregory Henselman-Petrusek
Princeton University, Princeton, New Jersey, United States

Tyler Giallanza
Princeton University, Princeton, New Jersey, United States

Sebastian Musslick
Princeton University, Princeton, New Jersey, United States

Jonathan Cohen
Princeton University, Princeton, New Jersey, United States

Abstract
Artificial systems currently outperform humans in diverse computational domains, but none has achieved parity in speed and overall versatility of mastering novel tasks. A critical component to human success, in this regard, is the ability to redeploy and redirect data passed between cognitive subsystems (via abstract feature representations) in response to changing task demands. However, analyzing shared representations is difficult in neural systems with distributed non-linear coding. This work presents a simple but effective approach to this problem. In experiments, the proposed model robustly predicts behavior and performance of multitasking networks on natural language data (MNIST) using common deep network architectures. Consistent with existing theory in cognitive control, representation structure varies in response to (a) environmental pressures for representation sharing, (b) demands for parallel processing capacity, and (c) tolerance for crosstalk. Implications for geometric (dimension, curvature), functional (automaticity, generalizability, modularity), and applied aspects of representation learning are discussed.