Bounded hypothesis testing underlying human learning of probabilistic rules

Jianbo Chen
Peking University, Beijing, China

Tianyuan Teng
Max Planck Institute for Biological Cybernetics, Tuebingen, Germany

Yuxiu SHAO
Beijing Normal University, Beijing, China

Hang Zhang
Peking University, Beijing, China

Abstract

We investigated human learning of probabilistic rules in experiments using a 2-by-2 feature space. Despite the seemingly minimal complexity, participants struggled to learn nonlinear XOR rules (where outcomes depend on cue matchings) but rapidly mastered linear rules. This difficulty persisted even when explicit probes revealed the possible rules, indicating constraints in hypothesis testing. To explain these behavioral patterns, we propose a hypothesis diffusion model where learning arises from evidence-driven transitions between hypotheses in a sparsely connected network. The model outperformed reinforcement learning alternatives and generalized across different rules. To further understand the origin of the learning difficulty, we trained low-rank recurrent neural networks and found that networks with limited capacity (rank 3) failed to learn XOR rules when trained in biased environments, mirroring human performance. In conclusion, human rule learning may rely on structured hypothesis exploration, with learning biases potentially emerging from adaptations to environmental demands under computational constraints.

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