LEAPS OF INSIGHT IN HUMANS LEARNING A COMPLEX TASK

Gautam Agarwal[1], Dongrui Deng[2], Mani Hamidi[3], Tiago Quendera[4], Mattia Bergomi[5], & Zachary Mainen[4]

[1] W.M. Keck Science Department at the Claremont Colleges
[2] Xi'an Jiaotong University
[3] Machine Cognition Lab, University of Tubingen
[4] Champalimaud Centre for the Unknown
[5] Veos Digital

Grant 360/18

Background: Learning a complex skill requires searching a potentially enormous search space. While reinforcement learning (RL) algorithms can approach human levels of performance in complex tasks, they require much more training than humans. This may be because only humans can infer and apply generalizable principles from limited experiences. Importantly, the statistics that underlie this learning process are both poorly understood and hard to investigate in the large state spaces found in most complex tasks.

Aims: Our goal is to design a cognitive task that is simple enough to study using the tools of psychophysics and RL, but complex enough to confront subjects with a large search space that compels them to use efficient solutions. We would like to define the statistical principles that shape these efficient solutions in order to understand complex skill learning in humans and inspire better artificial intelligence (AI) algorithms.

Method: We designed a task in which the subject must learn a sequence of actions that allows them to collect a reward. They receive minimal verbal instruction and can potentially attempt thousands of viable action sequences, only a few of which meet the criterion of success. We launched the game as a smartphone-based app (hexxed.io), allowing us to collect data from almost 10,000 participants attempting to solve the same problem. As a point of comparison, we run model-free RL agents known as Deep Q-Learning Networks (DQNs) on the same task.

Preliminary results: We find that unlike DQNs, 1) humans sample a highly restricted subset of the policy space, suggesting they draw from a "universal prior"; 2) humans repeatedly sample the same policy repeatedly, whether or not it is rewarding, suggesting they accumulate evidence against a theory before discarding it; 3) humans arrive at the optimal policy suddenly and unpredictably, reminiscent of the phenomenon of *epiphany*. We believe this reflects a "top-down" learning process in which people learn by proposing explanatory theories which they refine or replace only upon collecting sufficient evidence to the contrary.

Keywords: Decision making; Reinforcement learning; Complex skill; Video game; Intelligence; Epiphany; Insight

Email contact: gagarwal@kecksci.claremont.edu