SfN: Neuroscience 2008. Washington DC, November, 2008

A neurocomputational model for decreased harm avoidance in addicts

Amir Dezfouli Payam Piray Mohammad Mehdi Kramati
Hamed Ekhtiari Caro Lucas Azarakhsh Mokri

Abstract

Lack of harm avoidance is one of the most important decision making deficits in addicts. In previous computational models of addiction, this property is explained by high incentive value of drug of abuse which grows unbounded with consumption and compensates punishment of harmful consequences. But this theory is not confirmed by some animal models of drug self-administration which report unchanged incentive value of the drug before and after subjects show compulsive drug seeking. In this study we are to propose a computational model based on theoretical models of dopamine system and Bayesian Q-learning to shed light on the basis of this phenomenon. The agent learns values of actions in Bayesian fashion. Its action selection is guided by both expected state-action values and expected gain from exploring an action. The later indicates that the agent takes an action with low expected value if it predicts that the lost utility can be compensated by future gains of obtained information after executing the action (in the form of action selection policy improvement). We modeled rewarding pattern of natural rewards and cocaine in the manner that they have the same maximum reward value. Drug induced transient increase in dopamine was modeled by impulse rewards those vanish as the expected value saturates at its maximum level. The agent was simulated under situation where it should choose between avoiding harm(freezing) and pressing seeking lever which causes receiving a punishment followed by reward delivery. The results show in the case of natural reward, the agent chooses freezing, but in the case of drug, although expected value of pressing seeking lever dose not differ form natural rewards, the agent chooses drug seeking. Tracking agent's action selection policy revealed that such risk-taking behavior is because of high predicted gain of policy improvement after experiencing cocaine intake. Roughly speaking, the addicts chooses drug because it unrealistically predicts that after consumption, estimated value of drug taking will be updated to the value higher than freezing. It means that it hopes the punishment to be removed in the future. So, it keeps drug taking against harmful consequences with the hope of gaining utility more than harm avoidance. Our model proposes a biologically-inspired computational base for compulsive drug seeking and risky decision-making in addicts.