

Optimal learning: a route to depression?

Quentin Huys and Peter Dayan

Gatsby Computational Neuroscience Unit, UCL

Despite their epidemiological importance, psychiatric diseases are only poorly understood. Nevertheless, judging by the efficacy of pharmacological treatments, most appear to involve characteristic dysfunctions of the dopaminergic (DA), serotonergic (5HT) and/or noradrenergic neuromodulatory systems that are the foci of intense computational neuroscience studies. We seek to use this tie to provide a link from normative neuromodulatory models to psychiatric dysfunction, here in particular in the context of extensively validated animal models of depression.

The main clinical characteristic of depression is anhedonia – a disability to perceive or act upon rewards. This would seem to fit well into a reinforcement learning (RL) framework, providing a possible functional tie to DA and 5HT. However, key animal models observed to lead to anhedonia are learned helplessness (LH) and chronic mild stress, and LH in particular is taken to argue for the importance of the perception (and systematic generalizability) of controllability – a concept alien to standard RL treatments.

We study three facets of RL models of depression. First, we show that it is possible to capture key aspects of the LH data in a habit-based (average-reinforcement learning) model, independent of any explicit notion of controllability. This works by generalizing to other contexts any irrelevance found for action choice in one context. Human behavioral data support this idea, which also has wider application to such things as understanding the emotional regulation of pain.

Second, we construct an explicit, hierarchical Bayesian account of control. We consider as readily controllable, environments in which single actions have low-entropy consequences, and for which each possible goal can be achieved by at least one action. We consider parameters describing these characteristics to be what generalize from one domain to the next. High-control priors lead to higher average reward predictions, as sometime rewarding actions are assumed to be anytime exploitable. Generalizations as to likely controllability determine optimal exploratory behavior. Only under high-control priors do actions differ substantially in terms of predicted outcome; low-control priors can give rise to the behavioral insensitivity to rewards mentioned above, which is measured in terms of action choice. We also consider the fraction of controllable reinforcement and find that this can account for the occurrence of anhedonia after both acute severe and chronic mild stress.

Finally, these formulations of controllability can be seen as placing simple priors on the decision trees of MDPs. Making inferences in large trees is computationally challenging, and standard practice is to prune branches that are either proveably or just likely to be inferior. We show that this pruning can explain some of the confusing effects of 5HT, which has been postulated to report negative values and inhibit actions that lead to negative outcomes; but decreases of which produce powerful relapses of depression. We appeal to a combination of Pavlovian and instrumental effects to account for the fact that *high* 5HT seems to produce high average reward predictions, and that *lowering* 5HT decreases the average reward and potentially induces or maintains a depressed state.

In sum, we show that learning about the relevance of reinforcers to environments and about levels of control accounts for major kinds of depressed animal and human behaviors; and that some of 5HT's contradictory functions are explained by a mapping onto RL. This provides a *normative* characterisation of putatively non-normative psychiatric effects.