A Connectionist Model of the Role of Dopamine in Incentive Salience and Temporal Difference learning

Angela J. Thurnham* (a.j.thurnham@herts.ac.uk), D. John Done (d.j.done@herts.ac.uk), Neil Davey* (n.davey@herts.ac.uk), Ray J. Frank* (r.j.frank@herts.ac.uk)**
School of Computer Science,* School of Psychology, **University of Hertfordshire,
College Lane, Hatfield, Hertfordshire. AL10 9AB England

Dopamine as a Prediction Error (PE) Signal in Reinforcement Learning

Dopamine is a neurotransmitter and disruption to this system has long been associated with the neuropsychiatric disorder, schizophrenia. Current theories of the effects of dopamine on behaviour focus on the role of dopamine in Reinforcement Learning, where organisms learn to organise their behaviour under the influence of goals, and expected future reward is believed to drive action selection.

Reward Prediction Hypothesis

Schultz and colleagues have demonstrated that dopamine neurons in the midbrain regions are particularly sensitive to rewarding events, and single cell recordings have identified a phasic dopamine burst of activity which is posited to be a reward prediction error (Schultz 2000).

Gating Hypothesis

It is suggested that phasic dopamine signals control the gating of new information to the prefrontal cortex, but at the same time dopamine mediates learning by improving the signal to noise ratio of inputs to receiving units (Cohen, Braver & Brown 2002)

Incentive Salience Hypothesis

Others believe that dopamine release assigns incentive value to objects or acts, transforming 'like' into 'want,' enabling reward-seeking behaviours (Berridge & Robinson 1998). However, this does not address the problem of *how* to choose actions.

Temporal Difference Learning

A computational theory has been proposed by Sutton (1988) of how actions are chosen, explaining that the dopamine phasic signal is used in two ways: (i) As a prediction error or learning signal used to create better estimates of future reward. (ii) Dopamine release also biases action selection towards situations that predict the best reward.

The Model

The symbolic models of McClure, Daw & Montague (2003) and Smith, Li, Becker & Kapur (2004) have gone some way to unite the psychological and computational theories mentioned above. McClure et al. model the acquisition process in Reinforcement Learning and link the ideas of Incentive Salience to Reward Prediction via Temporal Difference Learning, while Smith et al. model the role of dopamine in the generation of expected reward, independently of the acquisition process, linking Incentive Salience to Reward Prediction via the Gating Hypothesis. It has been suggested that dopamine generates both learning and gating effects and that both have similar parameters, namely the phasic dopamine signal (Montague, Hyman & Cohen 2004)

We will report on our current research into how the various theories map on to each other and how they relate to the seminal connectionist models of Servan-Schreiber, Printz & Cohen (1990) and Cohen and Servan-Schreiber (1992). We will also describe our connectionist implementation of the McClure et al. model, where the role of dopamine will be more biologically plausible than that of a symbolic model.

References

Berridge, K.C., & Robinson, T.E., (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews 28,* 309-369.

Cohen, J.D., & Servan-Schreiber, D., (1992). Context, Cortex, and Dopamine: A Connectionist Approach to Behavior and Biology in Schizophrenia. *Psychological Review 1,* 45-77.

Cohen, J.D., Braver, T.S., & Brown, J.W., (2002). Computational perspectives on dopamine function in prefrontal cortex. *Current Opinion In Neurobiology 12,* 223-229.

McClure, S.M., Daw, N.D., & Montague, P.R., (2003). A computational substrate for incentive salience. *Trends in Neuroscience 26(8),* 423-428.

Montague, P.R., Hyman, S.E., & Cohen, J.D., (2004) Computational roles for dopamine in behavioral control. *Nature 431,* 760-767.

Servan-Schreiber, D., Printz, H., & Cohen, J.D., (1990). A Network Model of Catecholamine Effects: Gain, Signal-to-Noise Ratio, and Behaviour. *Science 249,* 892-895.

Schultz, W., (2000). Multiple Reward Signals in the Brain. *Nature Reviews/ Neuroscience 1,* 199-207.

Smith, A., Li, M., & Kapur, S., (2004). A Model of Antipsychotic Action in Conditioned Avoidance: A Computational Approach. *Neuropsychopharmacology 29,* 1040-1049.

Sutton, R.S., (1988). Learning to Predict by the Methods of Temporal Differences. *Machine Learning 3,* 9-44.