Group 2: Project 3

1) Jepma & Nieuwenhuis (2011) have presented evidence from pupillary measurements that norepinephrine is related to the balance between exploration and exploitation. In particular, they suggest that uncertainty might be signaled by tonic norepinephrine (as measured by baseline pupil diameter), and that this signal predicts the shift from exploration to exploitation. A simple way to track this uncertainty (see Frank et al., 2009) is using a Bayesian model of the reward distribution for each option. Specifically, each option is associated with a unknown reward probability; the learner represents their belief about this reward probability using a beta(a,b) distribution. When a=b=1 at the beginning of learning, the distribution is uniform over all possible reward probabilities. Whenever a reward is observed, a' = a + 1, and whenever a reward is not observed, b' = b + 1. Assume that each option *i* is assigned a decision value according to:
$Q(i) = M(i) + \sigma * V(i)$
Where $M(i)=a/(a+b)$ is the expected reward probability, and
$V(i)=(a*b)/[(a+b)^2*(a+b+1)]$ is the variance of the beta distribution, playing the role of an "uncertainty bonus" weighted by parameter σ. Assume that actions are chosen according to a softmax policy: $P(i) = \exp\big(\beta * Q(i)\big) / \sum_j \exp\,(\beta * Q(j))$. Show how this model can account for the relationship between pupil diameter and exploration/exploitation under the assumption that baseline pupil diameter is proportional to the entropy of the distribution over actions.

2) What happens when you don't have the uncertainty bonus? How could you experimentally determine whether or not people use an uncertainty bonus?

3) Discuss how this model could be related to dopaminergic novelty bonuses (Kakade & Dayan, 2002).

References:

Cohen, J.D., McClure, S.M., & Yu, A. J. (2007). Should I stay or should I go? Exploration versus exploitation. *Philosophical Transactions of the Royal Society B, 362*, 933-942.

Frank, M. J., Doll, B. B., Oas-Terpstra, J., & Moreno, F. (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nature Neuroscience, 12*, 1062–1068.

Kakade, S. & Dayan, P. (2002). Dopamine: generalization and bonuses. *Neural Networks, 15*, 549–559.