COMPUTATIONAL ANALYSIS ON REWARD-MODULATED ACTIVITIES OF CAUDATE NEURONS.

H. Nakahara*1, T. Trappenberg1, O. Hikosaka2, R. Kawagoe2,
and Y. Takikawa2.
1Lab for Info. Synthesis, Brain Science Institute,
RIKEN, Wako, Saitama, 351-0198, Japan. 2Dept. of Physiol., Juntendo
Univ. Sch. of Med. Tokyo, 113-0033, Japan.

It has been proposed that the spiny neurons (Sps) in the striatum are involved in reinforcement learning (RL) based on the actor-critic scheme by use of a prediction error carried by nigro-striatal dopamine neurons (e.g. Houk et al., '95). This hypothesis, however, has not been examined quantitatively with respect to neural activities in the striatum.

Takikawa et al. (Soc Neurosci Abstr 23: 464, '97) reported a variety of reward-modulated Sp activities in the caudate (CD) in a memory-guided saccade task in which only one of four directions was rewarded in each session. Many Sps responded to the directional cue, but the response was modulated in two ways: 'reward-dependent', the preferred direction changed so that the response was greatest for the rewarded direction in each session; 'reward-conditional', the preferred direction did not change across sessions, but the response was enhanced for the rewarded direction.

The present study shows that these Sps activities can be described by our minimal model derived from a simple RL in the actor-critic scheme in which the 'critic' works to evaluate a current state while the 'actor' works to choose an optimal action given the state. In our model, the 'critic' showed activity similar to the reward-conditional Sp response, while the 'actor' showed activity similar to the reward-dependent Sp response. Our model also successfully simulated the time course of change in these Sp responses after the reward direction was changed. We also suggest that the 'reverse' reward-conditional Sp response found in the experiment (i.e., the least response for the rewarded direction) can be accounted for by our model, taking into account of the direct and indirect pathways in the BG. Finally, we show that further extensions of our model provide several specific predictions that are experimentally testable.

Supported by SPRF of RIKEN to HN, CREST of JST and JSPS Research for the Future Program to OH, RK, YT. and HN.