

# 633 slide05 correction

Yoonsuck Choe

March 7, 2011

In slide05, for the  $TD(\lambda)$  part, the following correction should be applied (Mitchell slides page 34 and 35). The fixed parts are indicated by the underbraces (.....).

- Incorrect

$$Q^\lambda(s_t, a_t) = r_t + \gamma \left[ (1 - \lambda) \max_a \hat{Q}(s_t, a_t) + \lambda Q^\lambda(s_{t+1}, a_{t+1}) \right]$$

- Correct

$$Q^\lambda(s_t, a_t) = r_t + \gamma \left[ (1 - \lambda) \max_a \underbrace{\hat{Q}(s_{t+1}, a)}_{(1)} + \lambda \underbrace{Q^\lambda(s_{t+1}, a_{t+1})}_{(2)} \right]$$

The slides have been updated on the course web page, accordingly.

Note that this equation also appears in the old exam, so fix it before you attempt to solve it.

The meaning of this is as follows. In the marked equation below, (3) is a *mix* of (1)  $\hat{Q}$  at  $s_{t+1}$  and (2) recursive  $Q^\lambda$  at  $s_{t+1}$ , using the  $\lambda$  parameter.

$$Q^\lambda(s_t, a_t) = r_t + \gamma \left[ \underbrace{(1 - \lambda) \max_a \hat{Q}(s_{t+1}, a) + \lambda Q^\lambda(s_{t+1}, a_{t+1})}_{(3)} \right]$$

\* Thanks to Gopal Kedia for bringing this to my attention.