

Multi-Head Q Estimation

willer

June 7, 2020

$$f(x) = Q_i(s, a), i = 1, 2, 3, \dots$$

$$\hat{Q}(s, a) = \frac{1}{n} \sum_i Q_i(s, a)$$

$$var(\hat{Q}) = \frac{var(Q_i(s, a))}{n}$$

Using Bagging method. Sample a batch from experience buffer. Train Q value-head. Repeatly sample and train another head.