

در اینجا می بینیم که چگونه می توانیم از یک مدل استفاده کنیم.

1) $Q(s, a)$ را از جدول یا شبکه عصبی دریافت کنید.
 2) $Q(s, a)$ را برای تمام a محاسبه کنید و $\max_a Q(s, a)$ را پیدا کنید.
 3) $Q(s, a)$ را برای تمام a محاسبه کنید و $\max_a Q(s, a)$ را پیدا کنید.

4) $Q(s, a)$ را برای تمام a محاسبه کنید و $\max_a Q(s, a)$ را پیدا کنید.
 5) $Q(s, a)$ را برای تمام a محاسبه کنید و $\max_a Q(s, a)$ را پیدا کنید.

6) $Q(s, a)$ را برای تمام a محاسبه کنید و $\max_a Q(s, a)$ را پیدا کنید.
 7) $Q(s, a)$ را برای تمام a محاسبه کنید و $\max_a Q(s, a)$ را پیدا کنید.

$$Q(s, a) \leftarrow (1 - \alpha) Q(s, a) + \alpha [Sample]$$

$$Sample = R(s, a, s') + \gamma \max_a Q(s', a)$$

این عبارت را می توانیم به صورت زیر بنویسیم:

سوالات ارزشمند

1) agent با محیط خود تعامل می کند و سعی می کند به هدف خود برسد.
 2) action: عملی که agent انجام می دهد.
 3) deterministic: محیطی که در آن هر عمل یک نتیجه مشخص دارد.

لواء سوات قبضہ ۲

$$Q(S_{i,a}) \leftarrow (1-\alpha)Q(S_{i,a}) + \alpha [\text{sample}]$$

$$\text{Sample} = R(\mathcal{B}, a, \mathcal{S}) + \underset{\substack{\uparrow \\ \text{max}}}{\gamma} \max_{a'} Q(\mathcal{S}, a')$$

\downarrow

disCont

Ques in state, how many state? ~ 100-200 states discont or 100
~ 0.2-0.3 superstates

