

Fixed Q-Targets

Definition: Fixed Q-Targets

- 예측 네트워크 $Q(s, a; \theta)$ 가 바뀌면 목표 값 $R + \gamma \max Q(s, a; \theta)$ 도 바뀌면서 불안정한 학습이 되는 현상을 방지하기 위해 별도의 목표 네트워크 $Q(s, a; \theta^-)$ 를 두고 즉시 업데이트가 아닌 일정 주기마다 업데이트 시키는 학습방법
- Let parameters θ^- be the set of weights used in the target, and θ be the weights that are being updated
- Slight change to computation of target value:
 - $(s, a, r, s') \sim \mathcal{D}$: sample an experience tuple from the dataset
 - Compute the target value by the target network θ^- : $r + \gamma \max_{a'} \hat{Q}(s', a'; \theta^-)$
 - Compute the action-value $\hat{Q}(s, a; \theta)$ by main Q-network θ

Target network w^- Main Q-network w

$$L(\theta) = \frac{\left(R + \gamma \max_{a'} Q(s', a'; w^-) - Q(s, a; w) \right)^2}{\text{Target value} \quad \text{Action-value}}$$