Initialize $\vec{\theta}$ arbitrarily

Repeat (for each episode):

  $\vec{e} = \vec{0}$

  $s, a \leftarrow$ initial state and action of episode

  $\mathcal{F}_a \leftarrow$ set of features present in $s, a$

  Repeat (for each step of episode):

    For all $i \in \mathcal{F}_a$: $e(i) \leftarrow e(i) + 1$

    Take action $a$, observe reward, $r$, and next state, $s$

    $\delta \leftarrow r - \sum_{i \in \mathcal{F}_a} \theta(i)$

    For all $b \in \mathcal{A}(s)$:

      $\mathcal{F}_b \leftarrow$ set of features present in $s, b$

      $Q_b \leftarrow \sum_{i \in \mathcal{F}_b} \theta(i)$

    $\delta \leftarrow \delta + \gamma \max_{b \in \mathcal{A}(s)} Q_b$

    $\vec{\theta} \leftarrow \vec{\theta} + \alpha \delta \vec{e}$

    With probability $1 - \varepsilon$:

      For all $b \in \mathcal{A}(s)$:

        $Q_b \leftarrow \sum_{i \in \mathcal{F}_b} \theta(i)$

      $a \leftarrow \arg\max_{b \in \mathcal{A}(s)} Q_b$

      $\vec{e} \leftarrow \gamma \lambda \vec{e}$

    else

      $a \leftarrow$ a random action $\in \mathcal{A}(s)$

      $\vec{e} \leftarrow 0$

  until $s$ is terminal