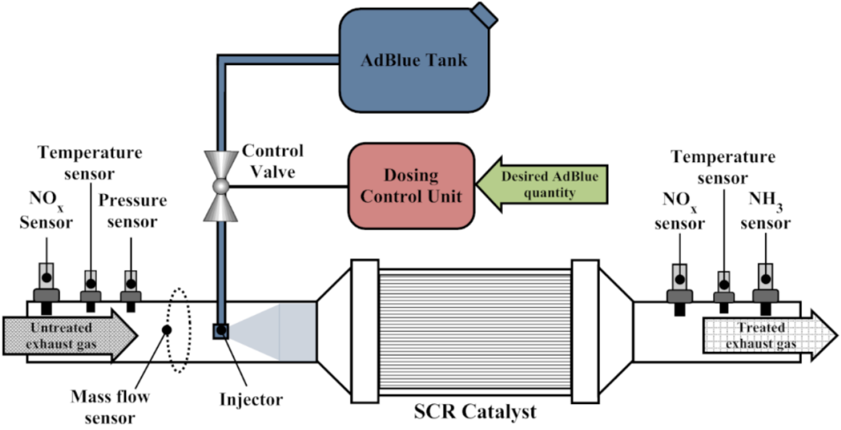
Selective Catalytic Reduction with Reinforcement Learning

June 12, 2019

The Current Process



Reinforcement Learning

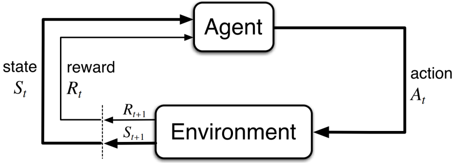


Figure: The agent-environment interaction in a Markov decision process [Sutton, R.S., Barto, A.G *Reinforcement Learning: An Introduction* (1998)]

Reinforcement Learning

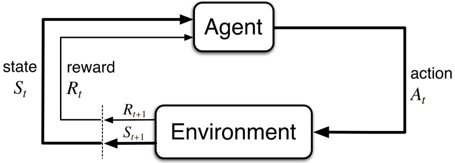


Figure: The agent-environment interaction in a Markov decision process [Sutton, R.S., Barto, A.G *Reinforcement Learning: An Introduction* (1998)]

Value & Q-Value Function

$$V^{\pi}(S) = \mathbb{E}_{\pi}\left[\sum_{i=0}^{T} \gamma^{i} R_{t+i} | S = S_t\right]$$

$$Q(S_t, A_t) = R_t(S_t, A_t) + \gamma \max_{A_{t+1}} Q(S_{t+1}, A_{t+1})$$

Reinforcement Learning

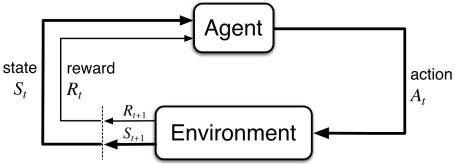


Figure: The agent-environment interaction in a Markov decision process [Sutton, R.S., Barto, A.G *Reinforcement Learning: An Introduction* (1998)]

Value & Q-Value Function

$$V^{\pi}(S) = \mathbb{E}_{\pi}\left[\sum_{i=0}^{T} \gamma^{i} R_{t+i} | S = S_t\right]$$

$$Q(S_t, A_t) = R_t(S_t, A_t) + \gamma \max_{A_{t+1}} Q(S_{t+1}, A_{t+1})$$

Minimise:

$$J = \left(Q(S_t, A_t) - R_t(S_t, A_t) - \gamma \max_{A_{t+1}} Q(S_{t+1}, A_{t+1})\right)^2$$

Reinforcement Learning

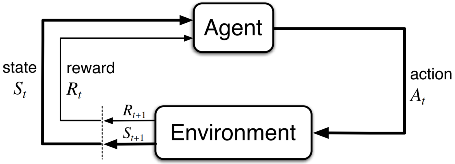


Figure: The agent-environment interaction in a Markov decision process [Sutton, R.S., Barto, A.G *Reinforcement Learning: An Introduction* (1998)]

Value & Q-Value Function

$$V^{\pi}(S) = \mathbb{E}_{\pi}\left[\sum_{i=0}^{T} \gamma^i R_{t+i} | S = S_t\right]$$

$$Q(S_t, A_t) = R_t(S_t, A_t) + \gamma \max_{A_{t+1}} Q(S_{t+1}, A_{t+1})$$

Minimise:

$$J = \left(Q(S_t, A_t) - R_t(S_t, A_t) - \gamma \max_{A_{t+1}} Q(S_{t+1}, A_{t+1})\right)^2$$

Update Equation:

$$Q(S_t, A_t) = Q(S_t, A_t) + \alpha\left(Q(S_t, A_t) - R_t(S_t, A_t) - \gamma \max_{A_{t+1}} Q(S_{t+1}, A_{t+1})\right)$$

Reinforcement Learning

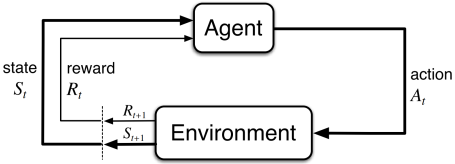


Figure: The agent-environment interaction in a Markov decision process [Sutton, R.S., Barto, A.G *Reinforcement Learning: An Introduction* (1998)]

Value & Q-Value Function

$$V^{\pi}(S) = \mathbb{E}_{\pi}\left[\sum_{i=0}^{T} \gamma^{i} R_{t+i} | S = S_t\right]$$

$$Q(S_t, A_t) = R_t(S_t, A_t) + \gamma \max_{A_{t+1}} Q(S_{t+1}, A_{t+1})$$

Minimise:

$$J = \left(Q(S_t, A_t) - R_t(S_t, A_t) - \gamma \max_{A_{t+1}} Q(S_{t+1}, A_{t+1})\right)^2$$

Update Equation:

$$Q(S_t, A_t) = Q(S_t, A_t) + \alpha\left(Q(S_t, A_t) - R_t(S_t, A_t) - \gamma \max_{A_{t+1}} Q(S_{t+1}, A_{t+1})\right)$$

Simplistic Reward Scheme:

$$R_t = \sum_{j=t}^{T+t} -\text{AdBlue}_j - c_1(\text{NH}_3)_j - c_2(\text{NO}_x)_j \quad \text{or} \quad R_t = \ldots?$$

Simpler Toy Problem

$$
\begin{aligned}
\frac{d}{dt}c_{NO,k} &= \frac{n}{V_c \cdot \varepsilon_g} \cdot \frac{\overset{*}{m}_{EG} \cdot R}{p_{EG} \cdot M_{EG}} \left(T_{EG,k-1} \cdot c_{NO,k-1} - T_{c,k} \cdot c_{NO,k}\right) + \\
&\quad + a_R \left(-4 \cdot r_{std,k} - 2 \cdot r_{fst,k} - r_{NO,g,k}\right) \\
\frac{d}{dt}c_{NO_2,k} &= \frac{n}{V_c \cdot \varepsilon_g} \cdot \frac{\overset{*}{m}_{EG} \cdot R}{p_{EG} \cdot M_{EG}} \left(T_{EG,k-1} \cdot c_{NO_2,k-1} - T_{c,k} \cdot c_{NO_2,k}\right) + \\
&\quad + a_R \left(-2 \cdot r_{fst,k} - 6 \cdot r_{slw,k} + r_{NO,g,k}\right) \\
\frac{d}{dt}c_{NH_3,k} &= \frac{n}{V_c \cdot \varepsilon_g} \cdot \frac{\overset{*}{m}_{EG} \cdot R}{p_{EG} \cdot M_{EG}} \left(T_{EG,k-1} \cdot c_{NH_3,k-1} - T_{c,k} \cdot c_{NH_3,k}\right) + \\
&\quad + a_R \left(-r_{ad,k} + r_{de,k} - 4 \cdot r_{ox,g,k}\right) \\
\frac{d}{dt}c_{O_2,k} &= \frac{n}{V_c \cdot \varepsilon_g} \cdot \frac{\overset{*}{m}_{EG} \cdot R}{p_{EG} \cdot M_{EG}} \left(T_{EG,k-1} \cdot c_{O_2,k-1} - T_{c,k} \cdot c_{O_2,k}\right) + \\
&\quad + a_R \left(-0.5 \cdot r_{NO,g,k}\right) \\
\frac{d}{dt}\theta_{NH_3,k} &= \frac{1}{\Theta_{NH_3}} \left(r_{ad,k} - r_{de,k} - 4 \cdot r_{std,k} - 4 \cdot r_{fst,k} - 8 \cdot r_{slw,k} - 4 \cdot r_{ox,k}\right) \\
\frac{d}{dt}T_{c,k} &= \frac{n}{m_c \cdot c_{p,c}} \left(\overset{*}{m}_{EG} \cdot c_{p,EG} \cdot \left(T_{EG,k-1} - T_{c,k}\right) + \alpha_c \cdot a_c \cdot \left(T_{Amb} - T_{c,k}\right)\right)
\end{aligned}
$$

$\Rightarrow$?