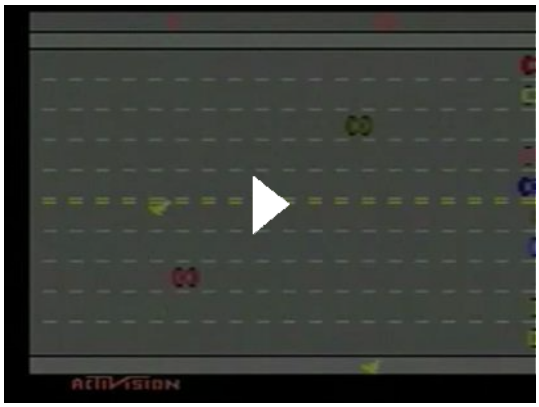
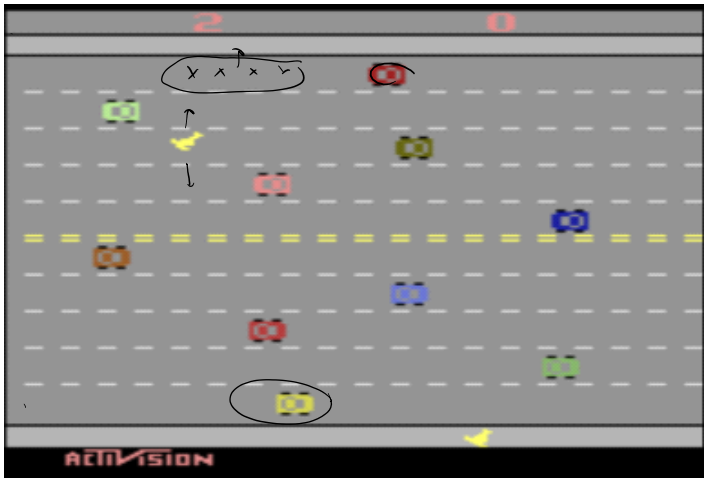
Q-Function Approximation

Learning to Play Freeway, using Reinforcement Learning





$Q(s,a) = f$

Features and representation:

- Feature function returning a feature vector: $f(s,a) = \begin{pmatrix} f_1(s,a) \\ \vdots \\ f_n(s,a) \end{pmatrix}$
- Weight vector: $n \times |A|$

Approximating Q(s,a)

$Q(s,a) = f_1(s,a) \cdot w_1^a + f_2(s,a) \cdot w_2^a \dots + f_n(s,a) \cdot w_n^a$

Q-learning Update:

execute a

→ for each feature i:

$w_i^a \leftarrow w_i^a + \alpha \left[r' + \gamma \max_{a'} Q(s',a') - Q(s,a) \right] f_i(s,a)$

same

$Q(s',a')$ for SARSA

Example (Freeway):
Assume Q(s,a) = 0 for all s,a

Update:

$w_i \leftarrow w_i + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s,a)] f_i(s,a)$

$\leftarrow 0 + 0.5 [10 + 0.9 \times 0] \frac{1}{1}$

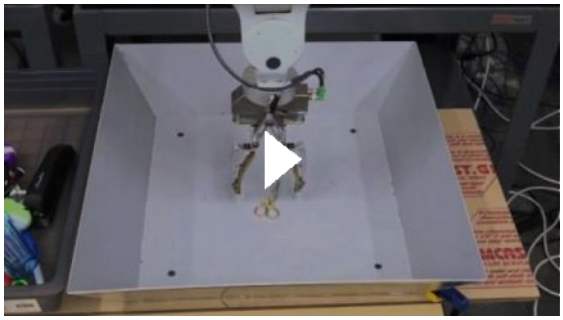
$w_i \leftarrow 5$

$\frac{1}{0} = \frac{1}{1}$

Deep Q learning:

$\theta \leftarrow \theta_i + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s,a)] \partial Q(s,a)/\partial \theta$

gradient

Deep Q-learning for robot grasping:

Learning Hand-Eye Coordination for Robotic Grasping