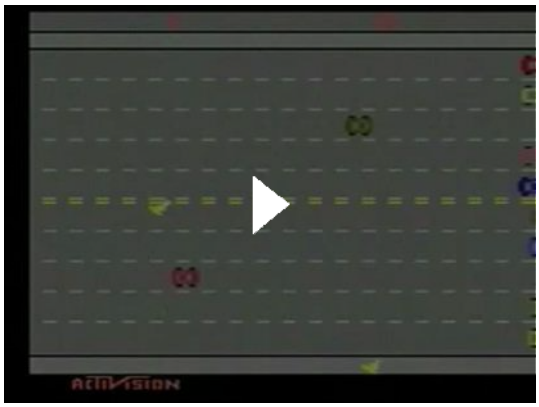
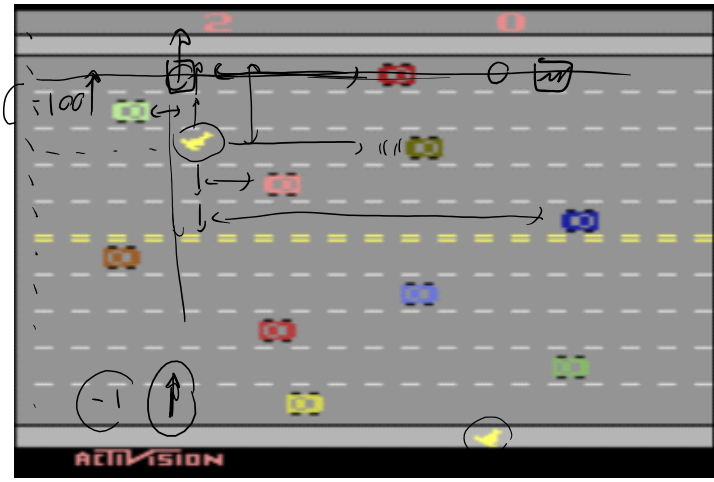
Q-Function Approximation

Friday, 21 September 2018 15:14 PM

[Learning to Play Freeway, using Reinforcement Learning]





$480^2 + 2^{480} = 3.12 \times 10^{44}$

(Distance1, Distance2, 1_{up}, 2_{up}, 1_{up}, 2Down)

Features and representation:

Feature vector: $f(s) = (f_1(s), \ldots, f_n(s))$

$f_1(s) = max\_D - row(s, c1)$

Weight vector: $w = \begin{pmatrix} w_1^a & w_1^b \\ w_2^a & \\ \vdots & \\ w_n^a & \end{pmatrix}$ $n \times |A|$

Approximating Q(s,a)

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

$\sum_i^n f_i(s) \cdot w_i^a$

TD Update: for each i

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

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

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

$5 \leftarrow 0 + 0.5[10 + 0] \frac{1}{1}$

Deep Q learning:
$\vartheta_i \leftarrow \vartheta_i + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s, a)] \partial Q(s, a)/\partial \vartheta$