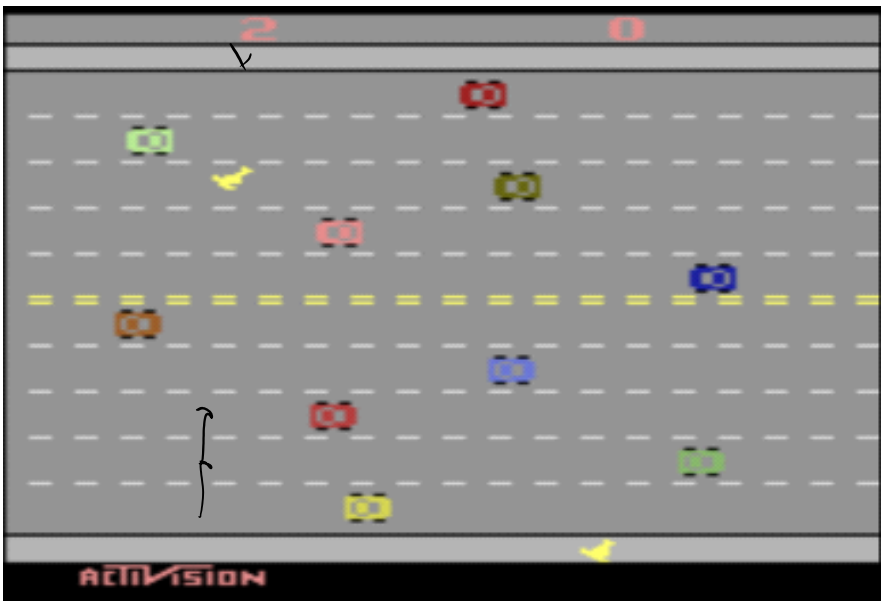
Reward Shaping

Friday, 21 September 2018 15:22 PM



Key idea: add some small additional reward for particular behaviour:

$Q(s,a) := Q(s,a) + \alpha[r + F(s,s') + \gamma \max a' Q(s',a') - Q(s,a)]$

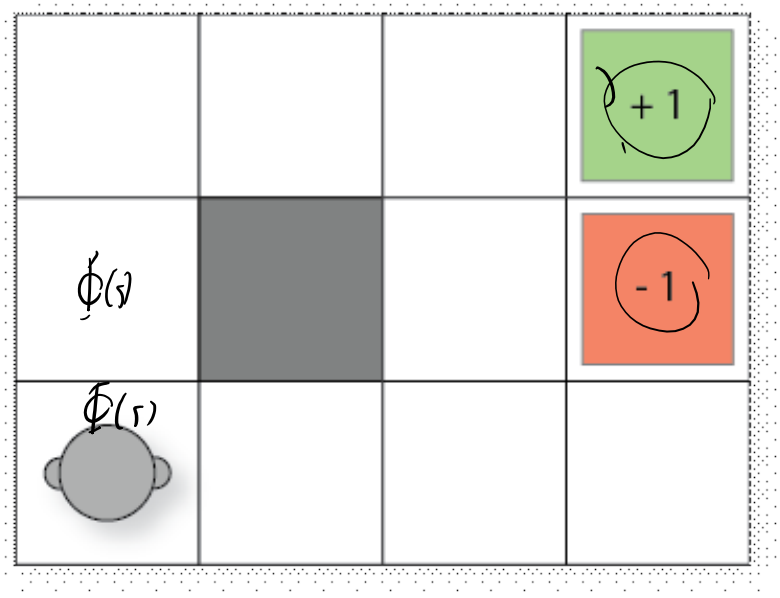
shaped reward

Potential-based Reward shaping:

$F(s,s') = \gamma \Phi(s') - \Phi(s)$

GridWorld example:

$\Phi(s) = -(|x(g) - x(s)| + |y(g) - y(s)|)$



F((1,2), (2,2)) = $\gamma \cdot \Phi(s') - \Phi(s)$

$= 0.9 \cdot \frac{1}{1} - \frac{1}{2}$

$= 0.4$

F((0,0), (0,1)) =

=

Alternative to reward shaping: Q-function initialisation