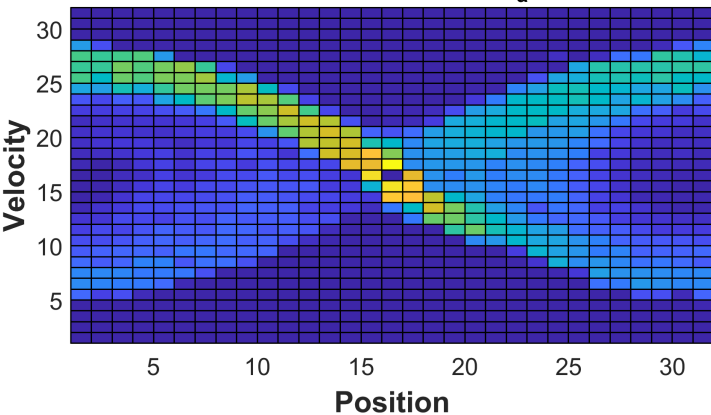


State-value function ($V = \max_a (Q(s, a))$)



Policy ($\arg\max_a (Q(s, a))$)

