

Parameter	Value	Description
γ	0.9	cumulative reward discount factor
α_π	0.001	actor learning rate
α_V	0.01	critic learning rate
momentum	0.9	stochastic gradient descent momentum
ϕ weight decay	0	L2 regularizer for critic parameters
θ weight decay	0.0005	L2 regularizer for actor parameters
minibatch size	32	tuples per stochastic gradient descent step
replay memory size	500000	number of the most recent tuples stored for future updates