Report

Learning Algorithm - MADDPG

This algorithm is best for problems where actions and states are in continuous space. It is an extension of the normal DDPG Algorithm but for 2 or more agents

Hyperparameters

BUFFER_SIZE = int(1e6) # replay buffer size
BATCH_SIZE = 128 # minibatch size
GAMMA = 0.95 # discount factor
TAU = 1e-3 # for soft update of target parameters
LR_ACTOR = 1e-4 # learning rate of the actor
LR_CRITIC = 1e-3 # learning rate of the critic
WEIGHT_DECAY = 0 # L2 weight decay
EPOCHS = 1 UPDATE_EVERY = 5

Model Architecture

3 Fully connected layers with 256, 128, 64 units respectively batch normalization on the input

Reward Plot