1. Parameters:

num\_epochs = 100

num\_episodes = 20 # number of episodes over which success rate is computed

episode\_length = 500

optimization\_steps = 40

actor\_learning\_rate=1e-3

critic\_learning\_rate=1e-3

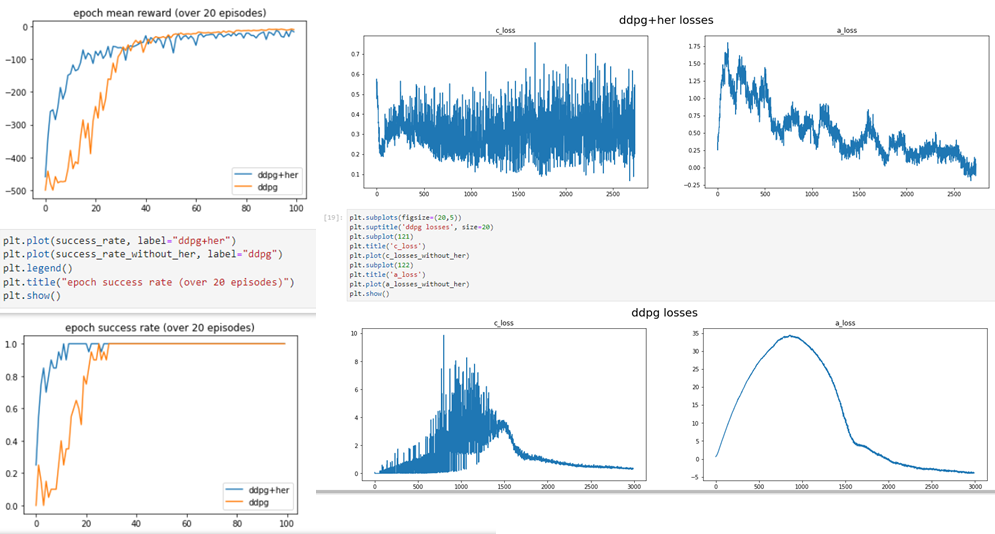
tau=0.1

Reward: -1 each step, +10 if touched detector

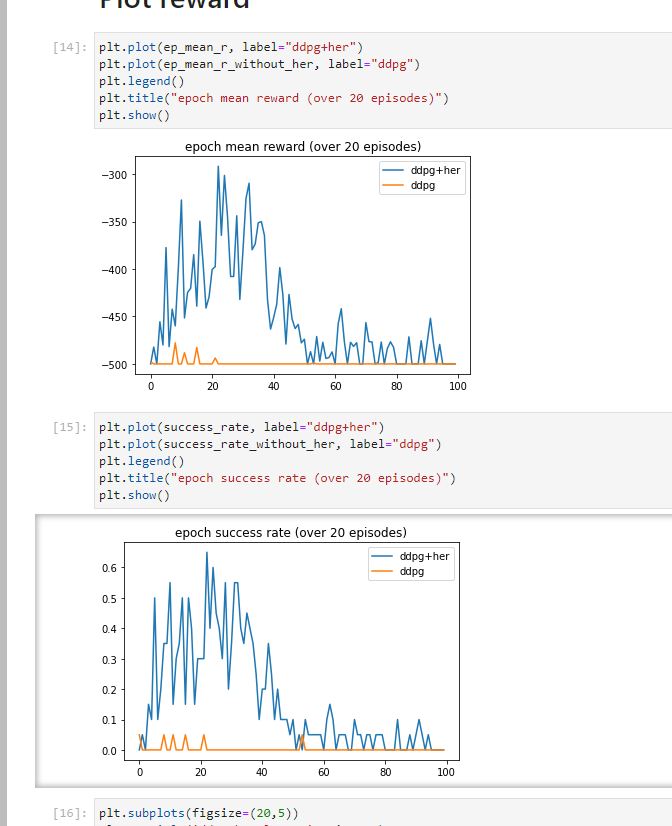
Saved models

HER: model/reward\_-1\_each\_step.ckpt

Without HER: model/reward\_-1\_each\_step\_without\_her.ckpt



1. Angles -120..120

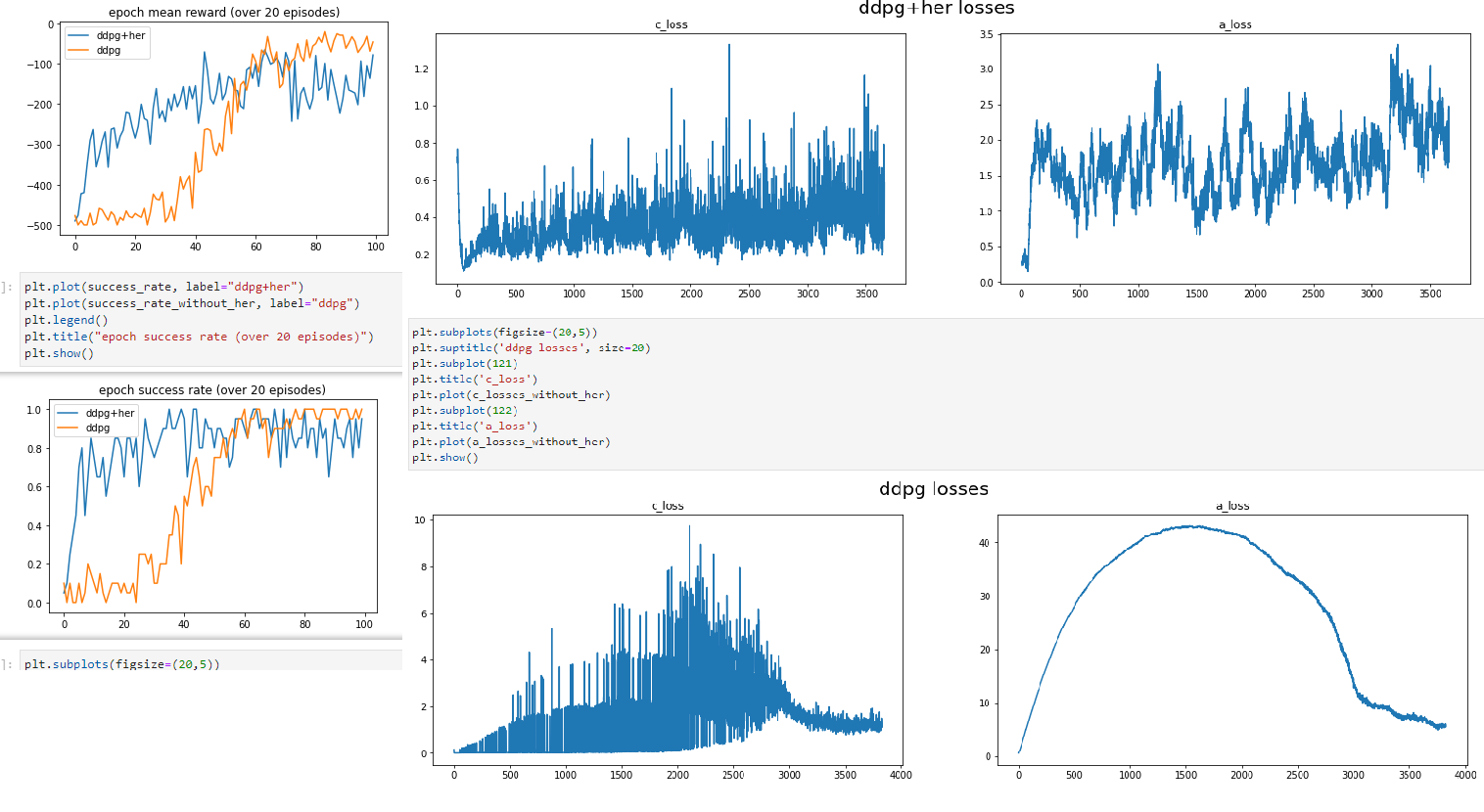


1. 3 actions – 2 angles for first hand, 1 – for second hand

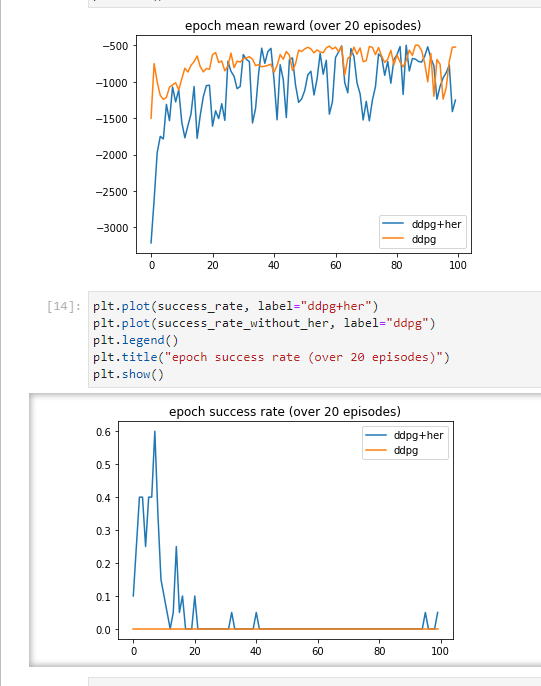
Models:

model/3actions\_not\_limited\_her.ckpt

model/3actions\_not\_limited\_without\_her.ckpt



1. -10 reward if out of angle range, 3 actions



1. Robohand with 3 arms, each arm have 1 angle

action\_high = 1, action\_low = -1

model/old\_hand\_hands3\_step3\_withoun\_her.ckpt

model/old\_hand\_hands3\_step3\_her.ckpt

