**实验记录**

**2023-5-25**

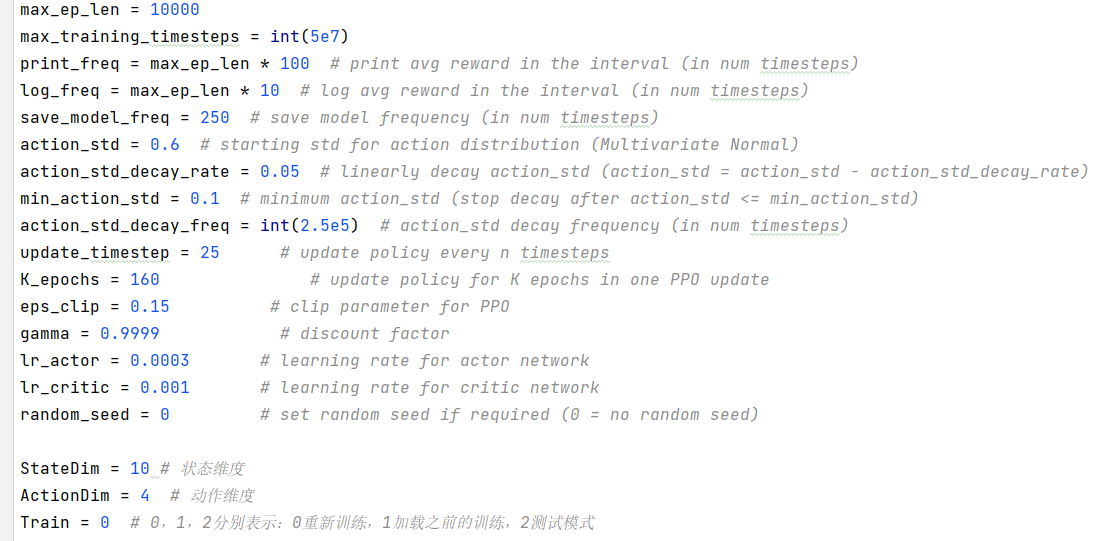
两机近距离追击任务

随机想定数量：1

想定设置：相对方位0度，相对朝向20度，其他默认

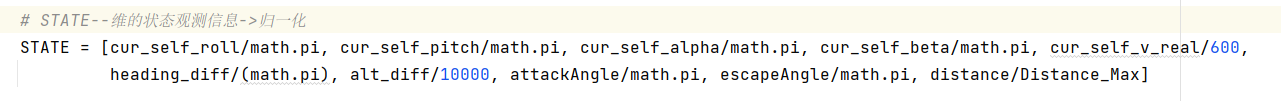
总结：此次更改可以收敛了，并且解决了不同朝向无法收敛的问题，测试并更改了朝向差、攻击角逃逸角相关的角度转换问题。并发现只要一个相对方位收敛了，就可以用到其他相似方位的想定里。

**1.超参数设置**：



**2.状态空间**（共10维）：

自身俯仰角、自身滚转角、alpha、bata、自身速度、两机航向差、两机高度差、攻击角、逃逸角、两机距离



**3.奖励设计**

1. 对局结束奖励：
2. 高度越界（小于1000米或大于10000米）

TotalReward = -50 + math.tanh((1000 - distance) / 10000)

1. 超时（300s）

TotalReward = -50 + 0.8\*math.tanh((1000 - distance) / 10000) + 0.2\*math.tanh(-math.fabs(attackAngle) / math.pi)

1. 达到追击距离（1000米）以及攻击角小于30度

TotalReward = 50 + 0.6\*math.exp(-CurTime / 300) + 0.2\*math.exp(-(math.fabs(attackAngle))/math.pi) + 0.2\*math.exp(-(math.fabs(escapeAngle))/math.pi)

1. 单步奖励：

dis\_reward = math.tanh((pre\_distance-distance)/10)

alt\_reward = math.tanh((pre\_alt\_diff-alt\_diff)/10)

attack\_reward = math.tanh((pre\_attackAngle-attackAngle)\*100)

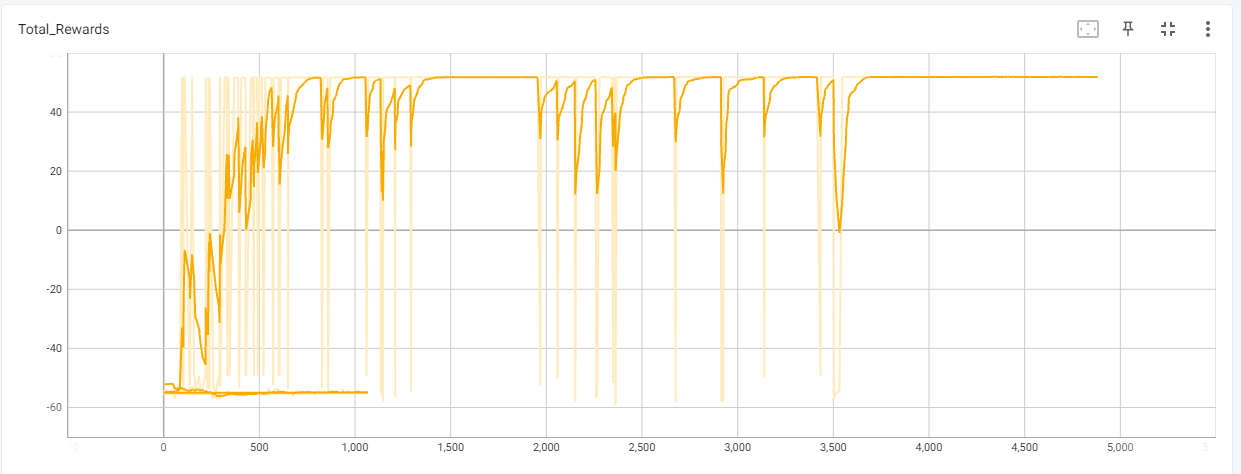
spd\_reward = math.tanh((distance/10000)\*(cur\_self\_v\_real-pre\_self\_v\_real))

if cur\_self\_alpha < -math.pi / 18 or cur\_self\_alpha > math.pi / 6:  
 alpha\_reward = -1

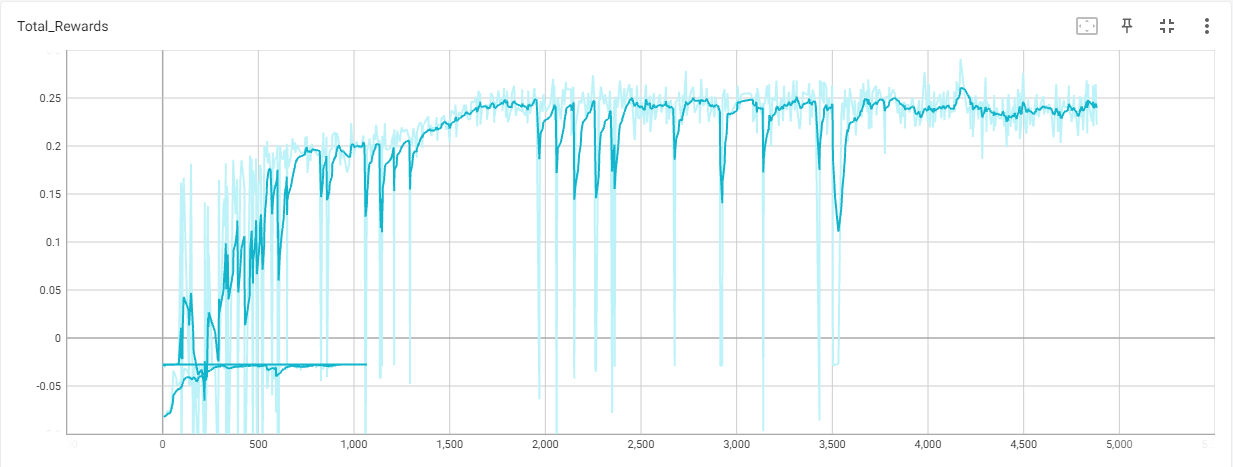
TotalReward = (0.5\*dis\_reward + 0.1\*alt\_reward + 0.3\*attack\_reward + 0.001\*spd\_reward + 0.1\*alpha\_reward)\*0.01

1. **实验结果**

每局总奖励和走势图



每局平均奖励走势图



第4000、4500、4700局单步奖励走势图

