

2020-10-14Configuration

Env Model	Differential wheeled model
Maximum linear V	1
Maximum angular V	3
reward in running	$(c_r \times (d_{t-1} - d_t) - k_1 \times (v_t - v_{t-1}) - k_2 \times (w_t - w_{t-1}))$
cr, K1, K2	[200, 0,0]
Terminal reward	(r_t)
rt, k3,k4	[50, 0,0]

Agent Model	PPO
Actor	<pre>nn.Linear(state_dim, 64), nn.ReLU(), nn.Linear(64, 32), nn.ReLU(), nn.Linear(32, 32), nn.ReLU(), nn.Linear(32,2), nn.Tanh()</pre>
Critic	<pre>self.state_block = nn.Sequential(nn.Linear(state_dim, 64), nn.ReLU()) self.outpu_block = nn.Sequential(nn.Linear(action_dim+64, 64), nn.ReLU(), nn.Linear(64, 32), nn.ReLU(), nn.Linear(32, 1), nn.Linear(1,1)) def forward(self, state, action): state_feature = self.state_block(state) merged = torch.cat((state_feature, action), dim=1) output = self.outpu_block(merged) return output</pre>

Result

Training Episodes	Result
100	cannot, flip
200	cannot, unstable
300	cannot
400	cannot
500	arrive at 1st target
600	finish 2 targets
700	out of boundary
800	out of boundary
900	out of boundary, almost flip
1000	
1500	out of boundary
1800	finished all avg 504 (rota, jump, dance)
1900	finished all avg 324
2000	cannot, (rotate at start point)

2020-10-11

Configuration

Env Model	Differential wheeled model
Maximum linear V	(1) 1.1
Maximum angular V	(3) 3.14
reward in running	$(c_r \times (d_{t-1} - d_t))$
cr, K1, K2	[200, 5,8]
Terminal reward	$(r_t - k_3 \times v_t - k_4 \times w_t)$
rt, k3,k4	[50, 5, 8]

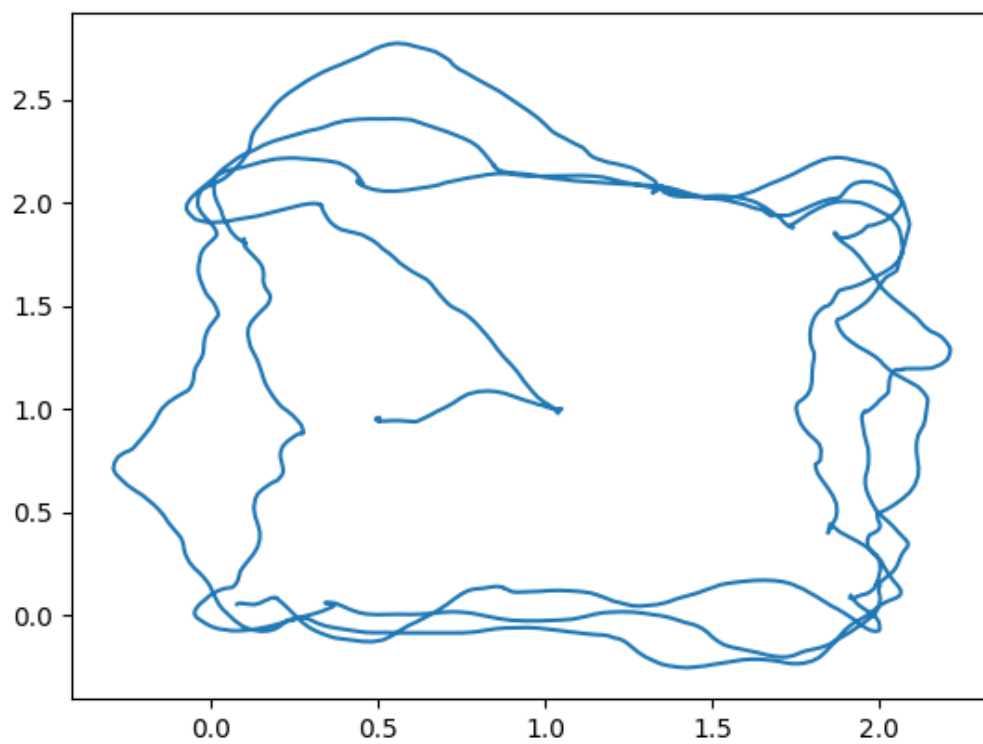
Agent Model	PPO
Actor	<pre> nn.Linear(state_dim, 64), nn.ReLU(), nn.Linear(64, 32), nn.ReLU(), nn.Linear(32, 32), nn.ReLU(), nn.Linear(32,2), nn.Tanh()</pre>
Critic	<pre> self.state_block = nn.Sequential(nn.Linear(state_dim, 64), nn.ReLU()) self.outpu_block = nn.Sequential(nn.Linear(action_dim+64, 64), nn.ReLU(), nn.Linear(64, 32), nn.ReLU(), nn.Linear(32, 1), nn.Linear(1,1)) def forward(self, state, action): state_feature = self.state_block(state) merged = torch.cat((state_feature, action), dim=1) output = self.outpu_block(merged) return output</pre>

Result

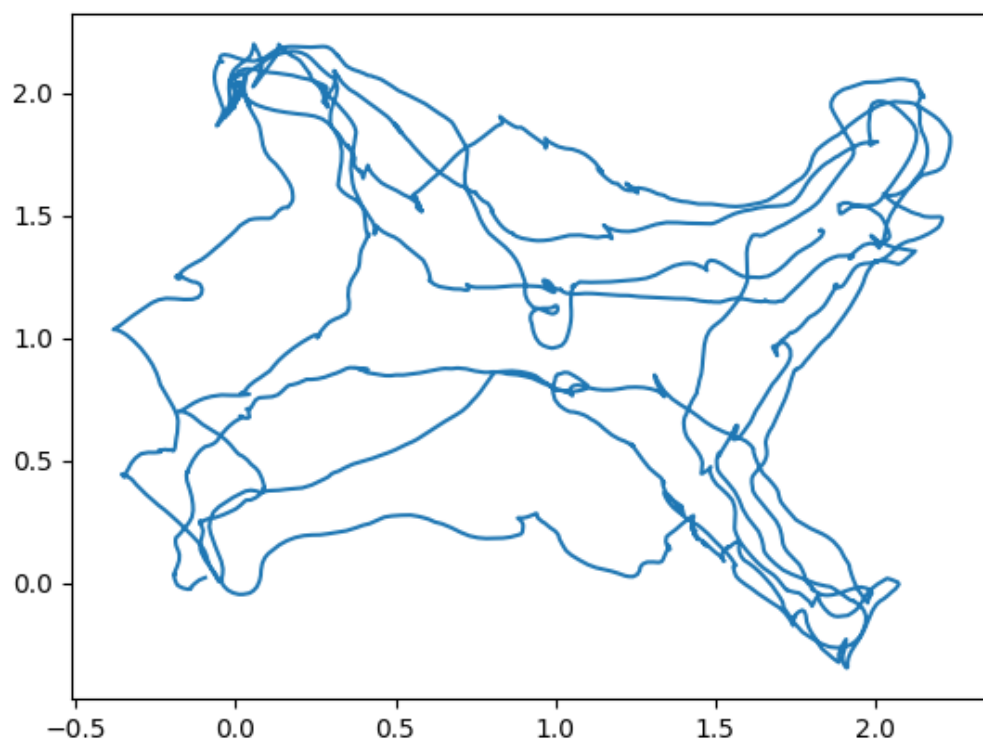
Training Episodes	Result
100	avg: 114
200	215.6
300	147
400	235
500	126.6
600	327
700	253.92307692307693
800	317
900	
1000	129
1100	flip
1300	cannot
1400	166
2100	446.6
2500	out of boundary
3000	695
4000	504
7400	cannot

Figure

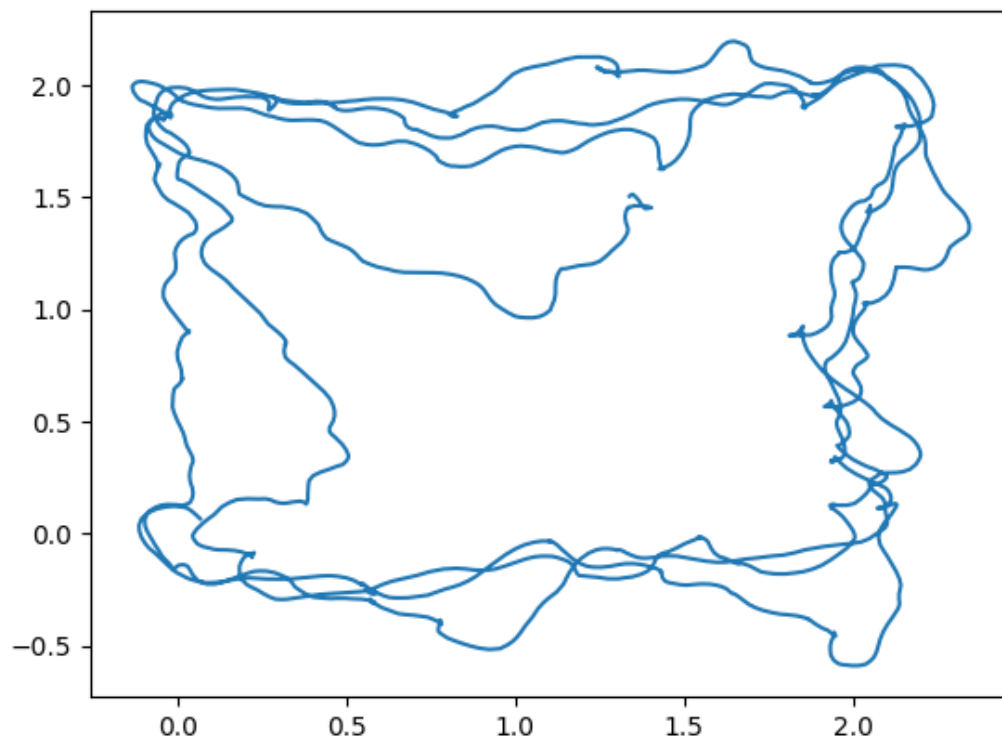
100 episodes



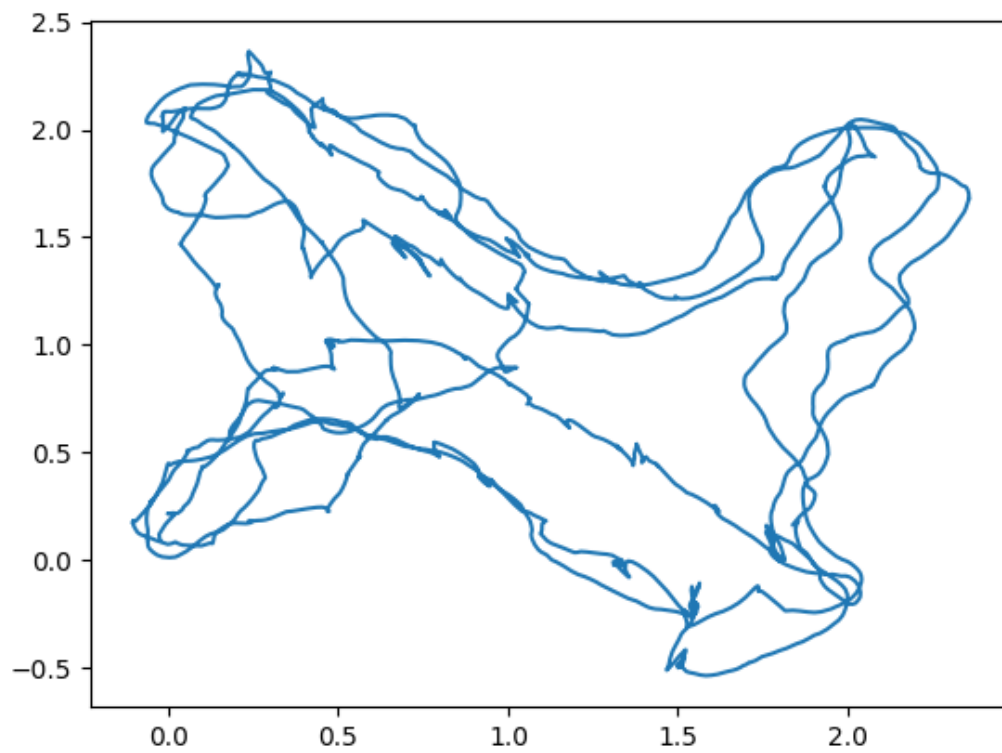
200 episodes



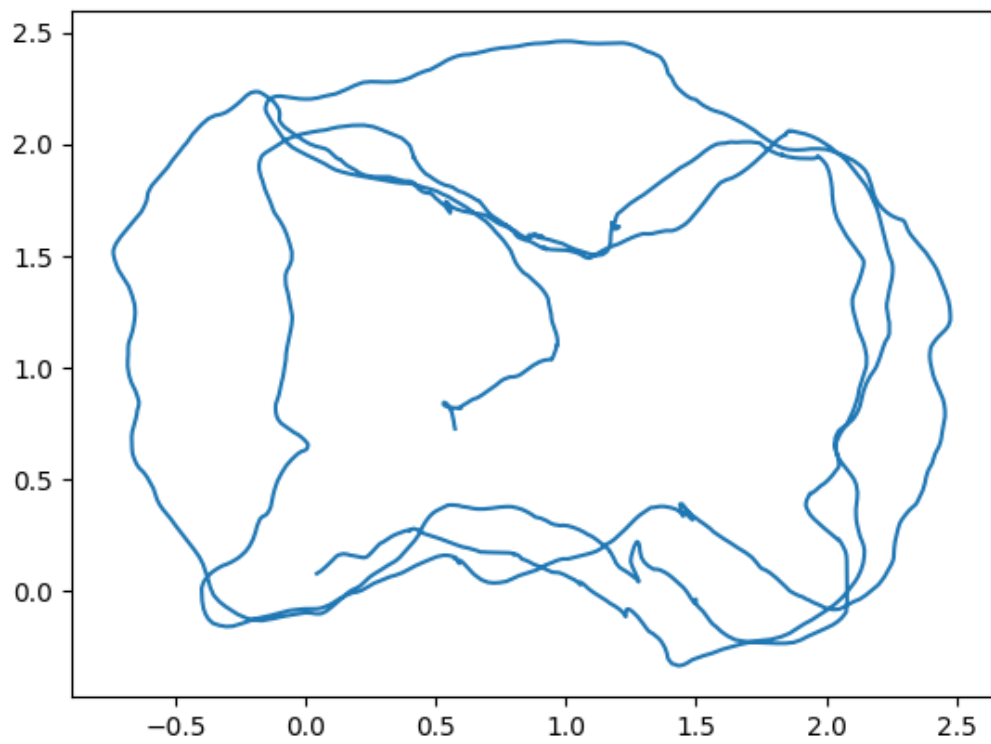
300 episodes



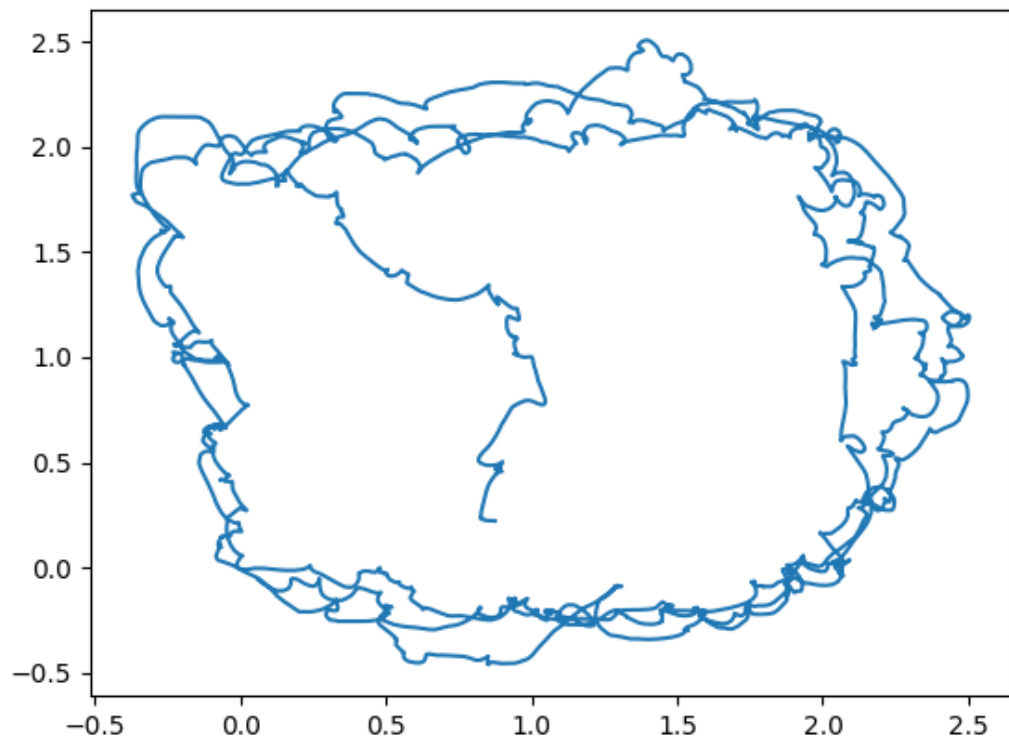
400 episodes



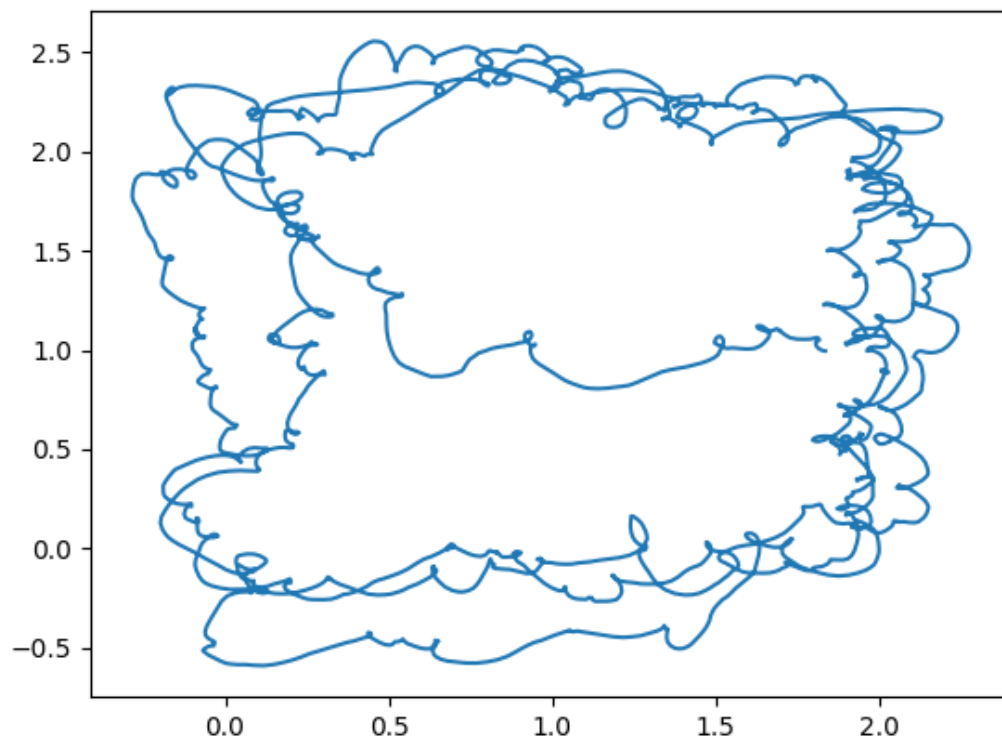
500 episodes



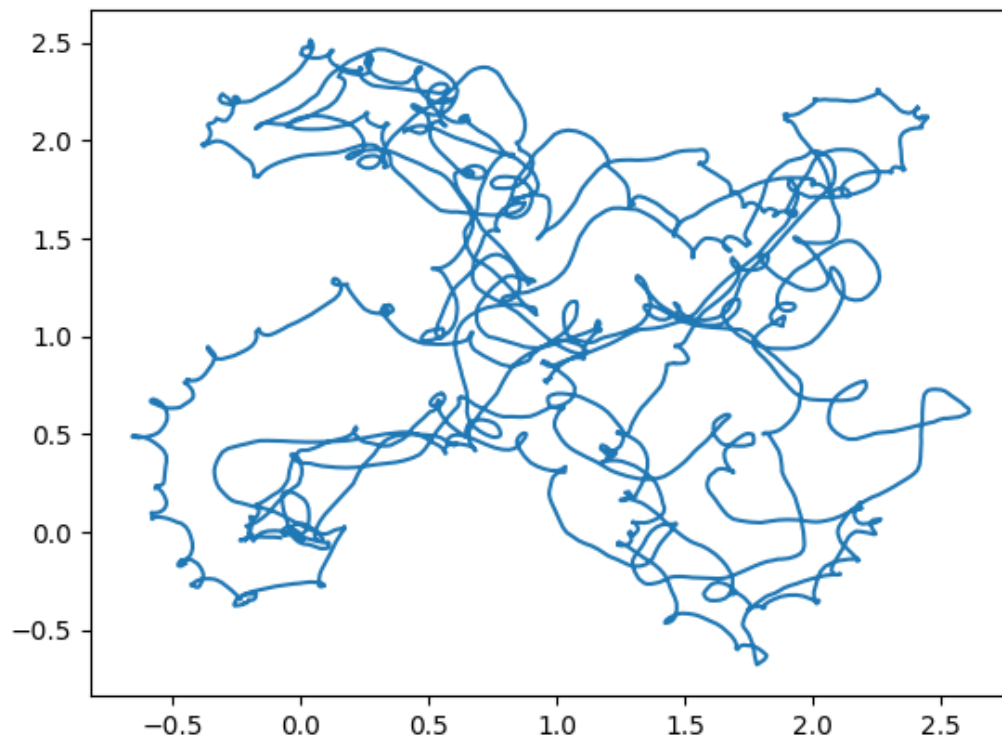
600 episodes



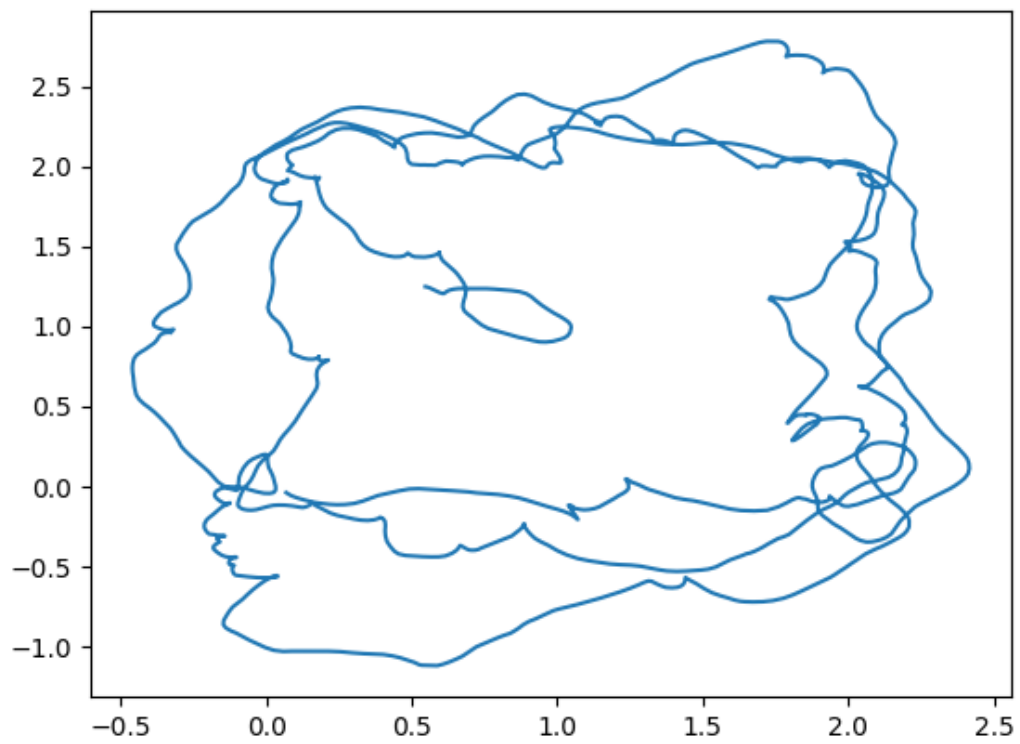
700 episodes



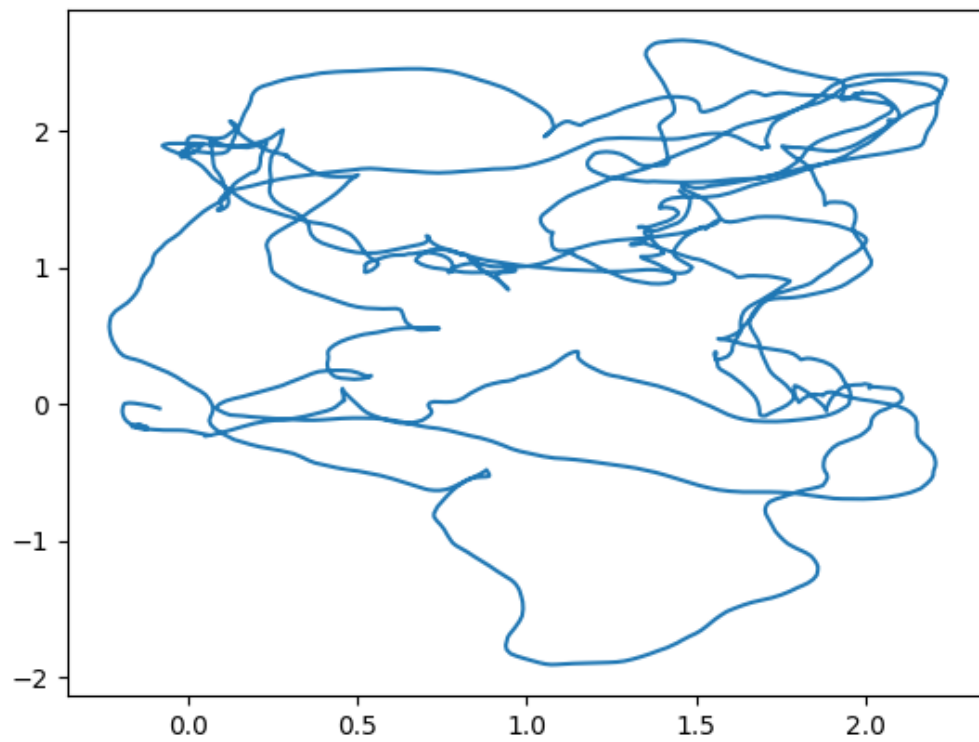
800 episodes



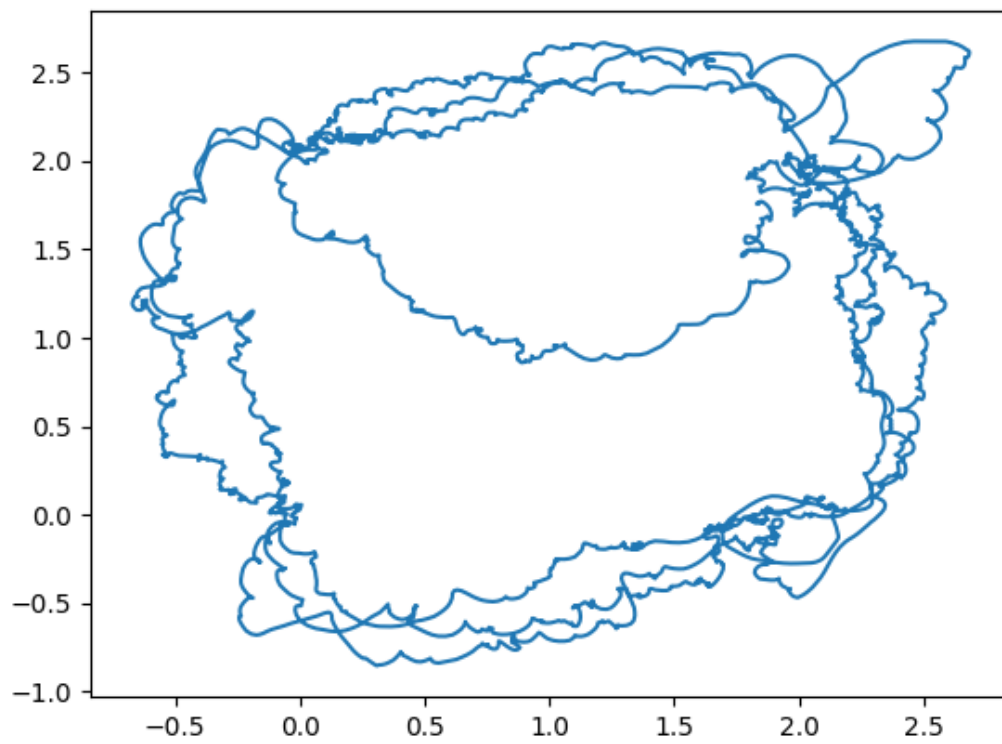
1000 episodes



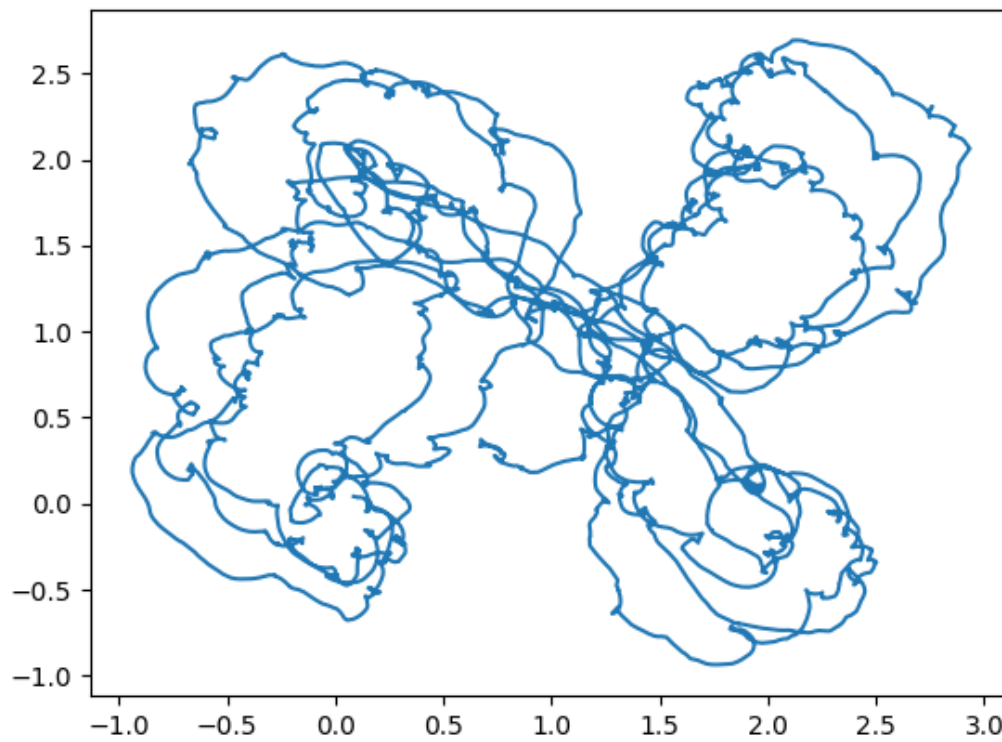
1400 episodes



2100 episodes



3000 episodes



4000 episodes

