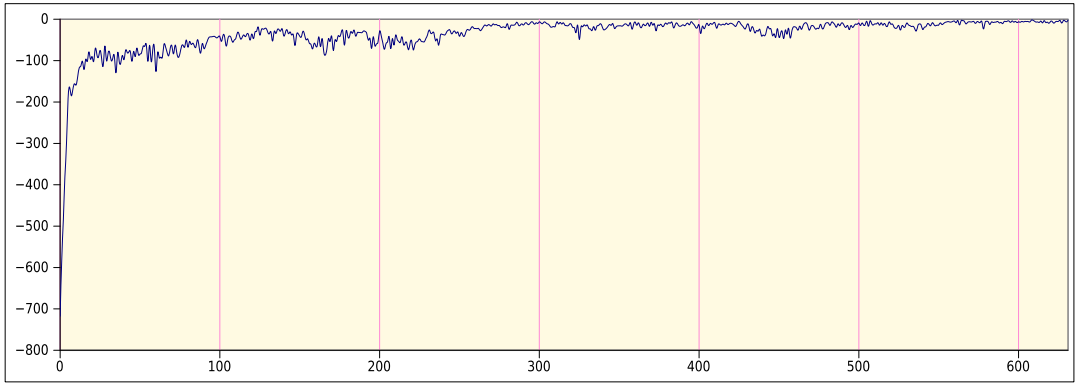
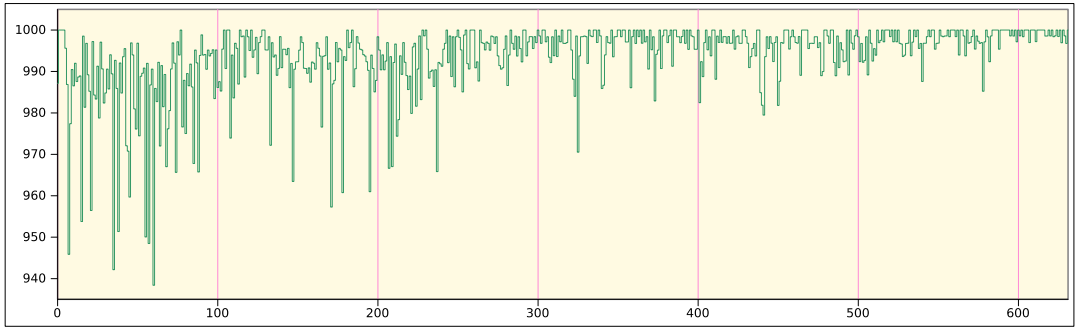
jokolaco

The gradient methods leads to very good results. This is the result of the first try of the implemented Algo without any improvements. It runs stable and smooth. Constant stepsize = 0.15, high evaluation loopcounts for the estimates.

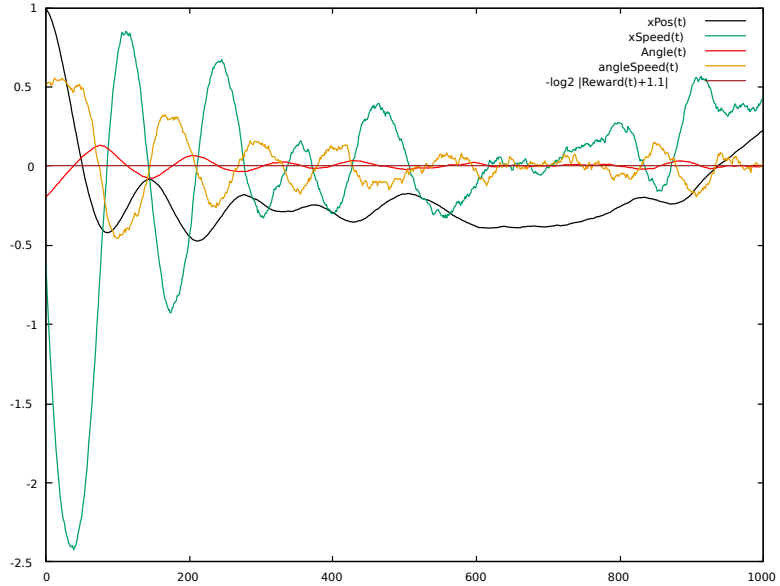


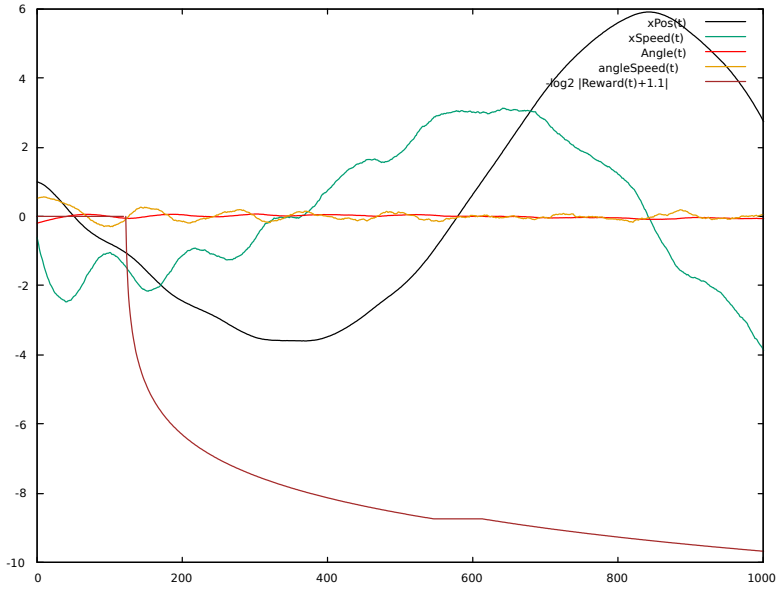
AvgReward



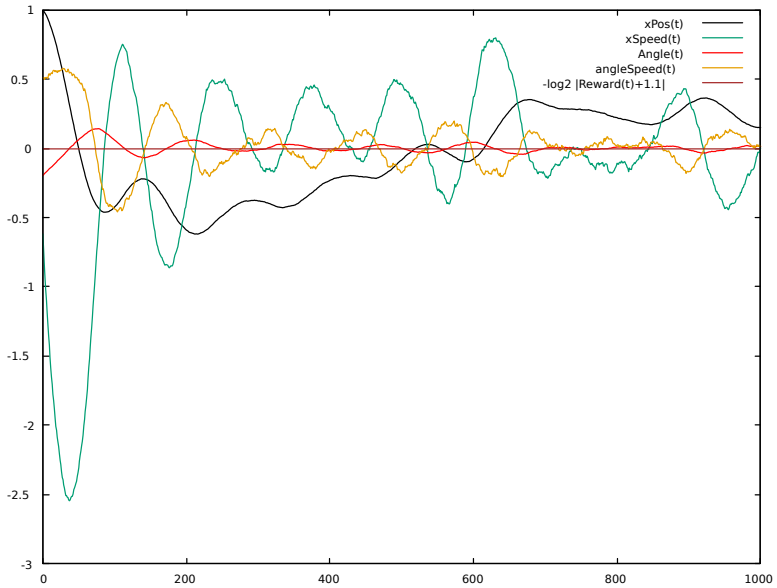
AvgEndtime

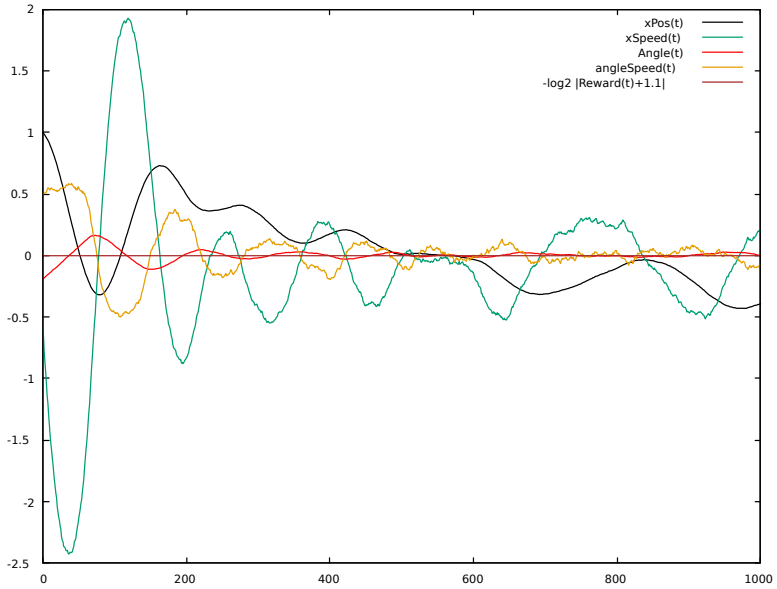
<u>BEFORE</u> Poliy k_1= 0.999416 k_2 = -0.340716 k_3 = 379.273005 k_4 = 4.172866 | Reward -3
Very Unstable, as in the other examples!



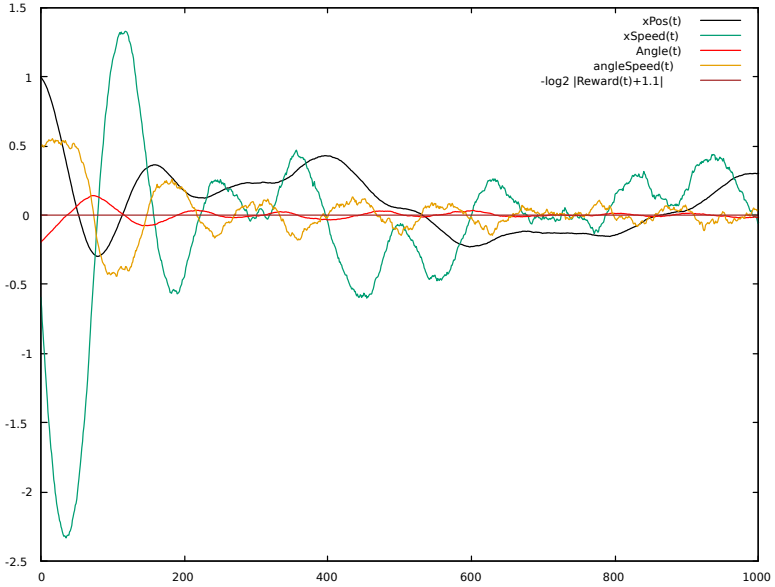


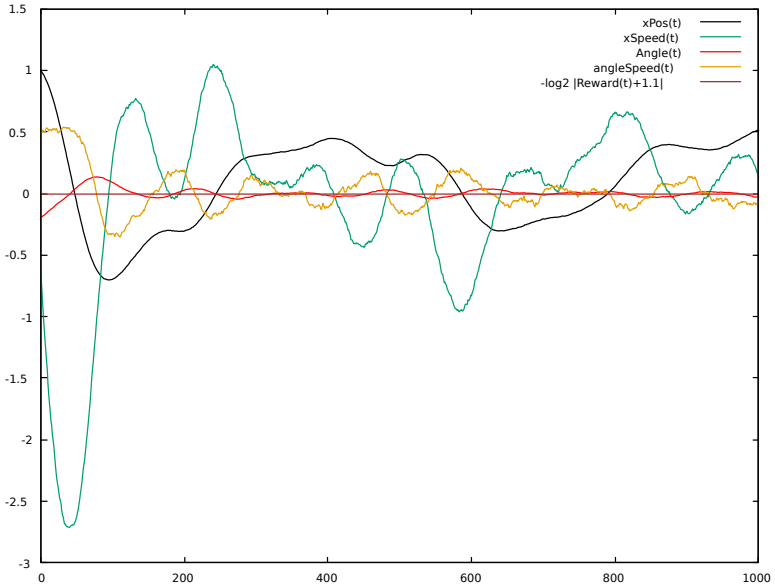
AFTER Policy with Gradient Ascent after 631 Steps and constant stepsize= 0.1 :
k[0]= 1.599416 k[1]= 2.059284 k[2]= 379.573005 k[3]= 11.072866





Stable in each run!





Stable in each run!