Homework 2 Submission
Pratap Bhanu Solanki Y9429
Paper Review of

Reinforcement learning to adjust Robot Movements to New Situations

Introduction: This paper focuses on modulating the elementary movement of Robot's actuators so that it can learn to adjust to new circumstances which are slightly different than the old training cases. Here Robot does not have to learn the motor parameters again. They introduced a reinforcement learning algorithm based on kernelized version of the reward weighted regression.

Methodology:

1. **Turning to Cost-regularized Kernel Regression(CrKR):** By inserting the reward weighted regression (RWR) solution into system output equation of basis function the reward weighted regression is turned into Cost–Regularized Kernel Regression.
2. The rewards are assumed to be positive and hence are transformed into cost by inverse relation. Now the cost is directly proportional to the distance from the desired optimal solution at a point.
3. **Meta-parameter learning by Reinforcement learning:** First of all this algorithm receives three inputs :
   - **a.** Motor primitives (of Each degree of freedom)
   - **b.** Initial example containing system states, meta-parameter and Cost
   - **c.** Scaling parameter

For Illustration of this algorithm an example of 2D dart throwing task is used. Here state 's' is height of the impact point. Meta-parameters are velocity and angle of the leaving dart.

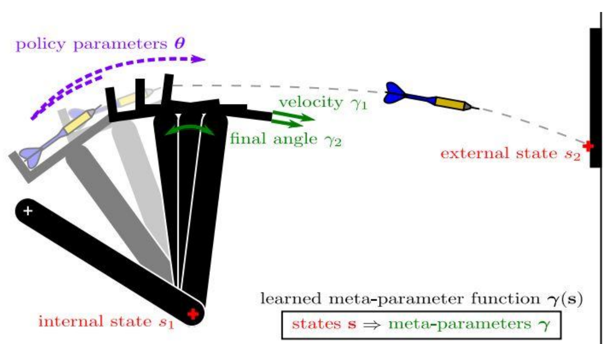

Figure: 2D Dart throwing task(Courtesy [1])

The cost here is error between desired goal and impact point and the launching velocity of dart. Initially meta-parameters are sampled from current state. Then from the outcome of trial cost is calculated and accordingly policy is updated. Similar process is followed later in each iteration.
The dart throwing is tested with a simulated robotic arm and also with a real Barett WAM.

Results Evaluation:

In simple cannon shooting they have bench marked the Reinforcement Learning by CrKR approach against Finite difference gradient estimator and reward-weighted regression.
In real complete framework, a complex dart throwing task called 'Around the Clock' was taken. Here also CrKR outperformed RWR which can be observed from figure 2.

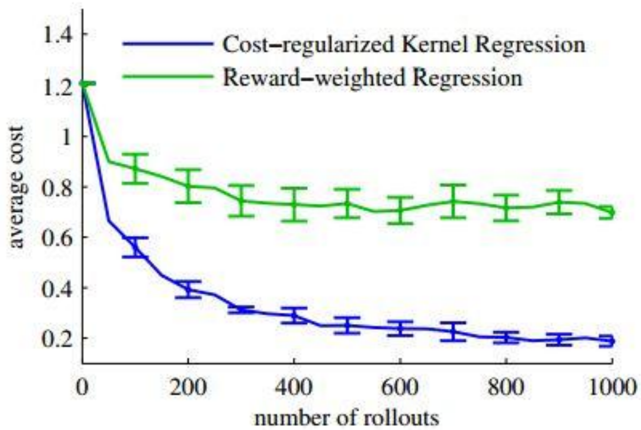

Figure: 2 Cost function for Dart throwing Task (Courtesy [1])

Table Tennis: The complete framework has also been tested for hitting a table tennis ball in the air. The meta-parameters are joints positions and velocities for all seven degrees of freedom plus a training parameter. These 15 parameters are learned and optimized using Gaussian kernel. At the beginning the robot misses 95% of the balls but finally it hits almost all balls (figure 3)

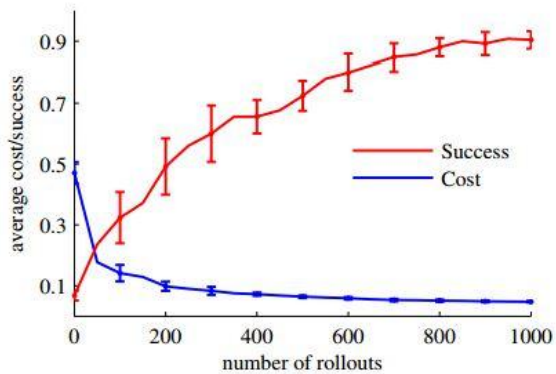

Figure 3: Cost function of Table Tennis task (Courtesy [1])

Conclusion: It can be seen form the results that the necessary mapping required at the time of new situation can be learn using a CrKR. This algorithm outperforms RWR and other preceding methods.

Reference:

1. Reinforcement Learning to adjust Robot Movements to new situations. Kober , Oztop, Peters Proceedings of the 22nd IJCAI Volume Three,2650--2655,2011.(Given paper)