Distributed Learning for the Decentralized Control of Articulated Mobile Robots

Yunfei Shi, Guillaume Sartoretti, William Paivine, Matthew Travers, Howie Choset

Objectives

- Improve the locomotion of articulated robots in unstructured environments.
- Explore the relationship between decentralized control and distributed learning.

Introduction

The locomotion of snake robots continues to be of interest to researchers because their high degree of actuation makes them appropriate for a variety of complex terrains and applications.

Figure 1: Snake robot in unstructured environments

we present a learning approach that leverages recent advances in distributed reinforcement learning to learn a decentralized control policy.

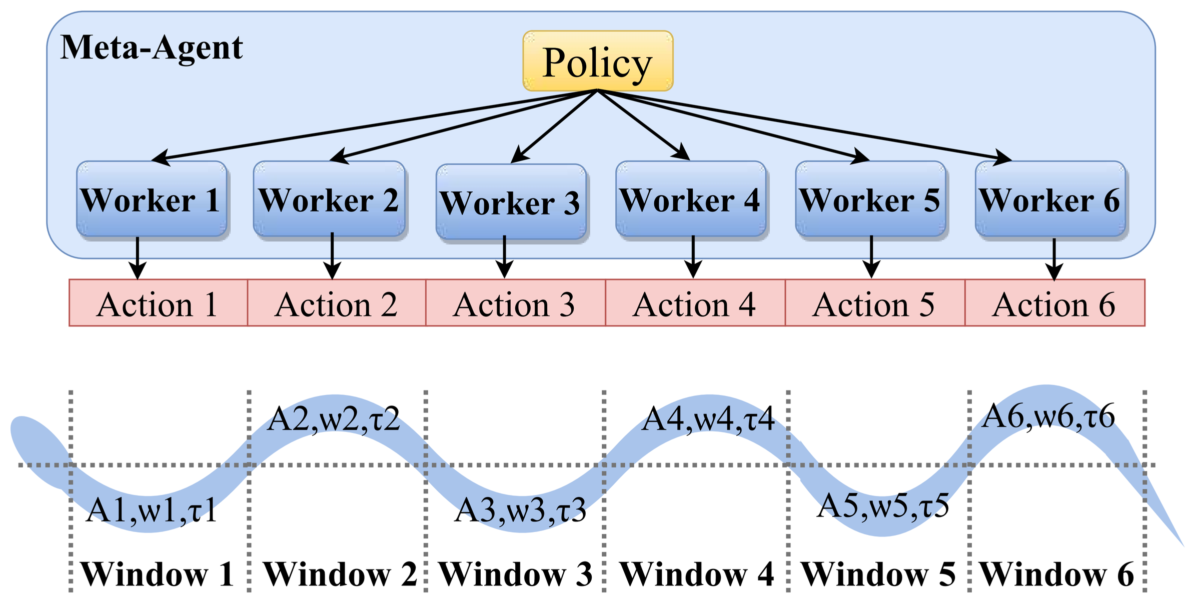

Figure 2: Analogy between A3C and the snake's control windows

Policy Representation

1. **State Space**
   - Modular time $\mu(t) = mod(t, T_S)/T_S$
   - Shape parameters $\beta(s,t)^T$
   - External torque in shape space $F(s,t)^T$
   - Nominal shape parameters β_0^T

$$s = \langle \mu(t),\ \beta(s,t)^T,\ F(s,t)^T,\ \beta_0^T \rangle. \quad (1)$$

2. **Action Space**
   Discrete increments in Amplitude and Spatial Frequency (SpF)
   - Amp. increment $a_A \in \{0, \pm\Delta_A\}$
   - SpF increment $a_\omega \in \{0, \pm\Delta_\omega\}$
3. **Actor-Critic NetWork**
   We design the Actor-Critic networks as 2 neural networks with weights Ψ_A, Ψ_C both using 4 fully connected layers to approximate the stochastic policy and the value function respectively.

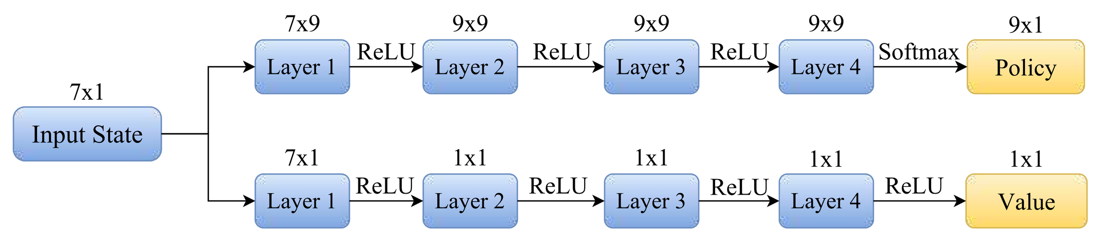

Figure 3: Actor-Critic Network Structure

Structure of the A3C meta-agent:

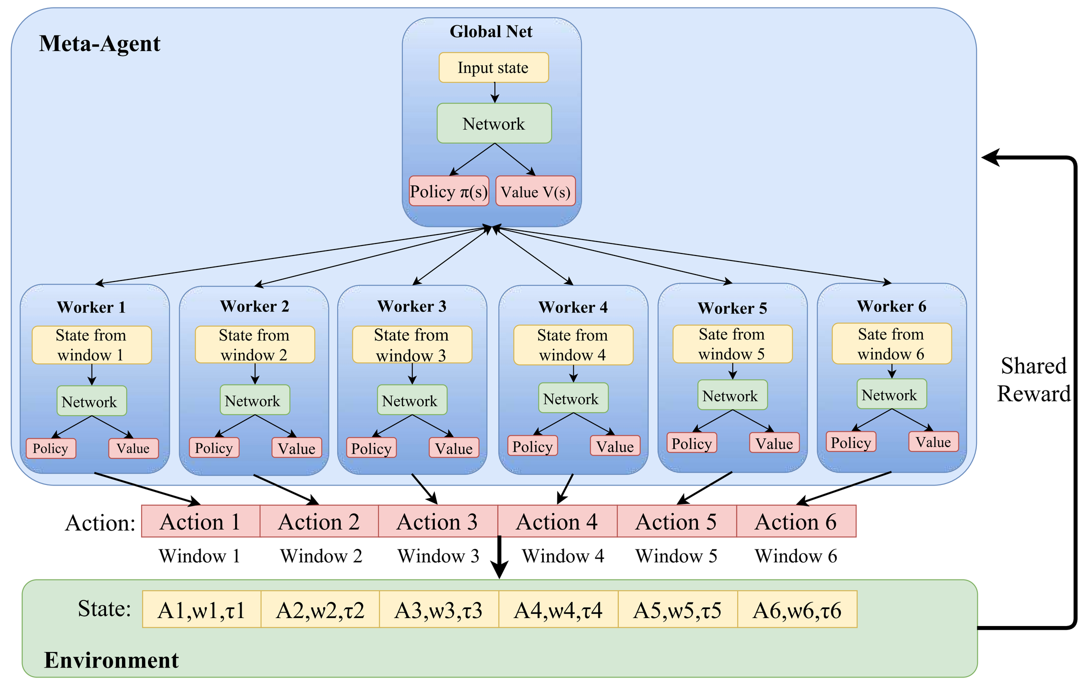

Figure 4: Structure of the A3C Meta-Agent

4. **Shared Reward**
   - Shared reward based on the instantaneous forward progression
   - Better for collaborative tasks (locomotion in our case)

$$r_t = \tanh\left(lambda_r \cdot \|X(t) - X_{0,i}\|_2\right), \quad (2)$$

Learning

The Learning data is collected on a pegboard:

Figure 5: Snake robot wearing a polyester sleeve traversing an unstructured peg array.

Rather than initializing the learning agent with randomly seeded parameters to sample the state-action space, we propose to use recorded actions of the current state-of-the-art compliant controller.

- 310 trials of compliant controller in Pegs
- 89 steps per episode to match 1/2 gait cycle
- Bootstrapping offline using A3C algorithm
- 6 agents are trained until 50,000 episodes
- the 6 windows correspond to the 6 agents

Experiment Validation

1. **Experiment Setup**
   - Platform: 16-module series elastic actuated snake robot
   - Only 8 planar modules were active
   - Tested 2 version of decentralized controller:
     1). Learning-based Controller
     2). Admittance Controller
     to adapt shape parameters
2. **Experiment Result**

Figure 6: Learning-based controller in peg array

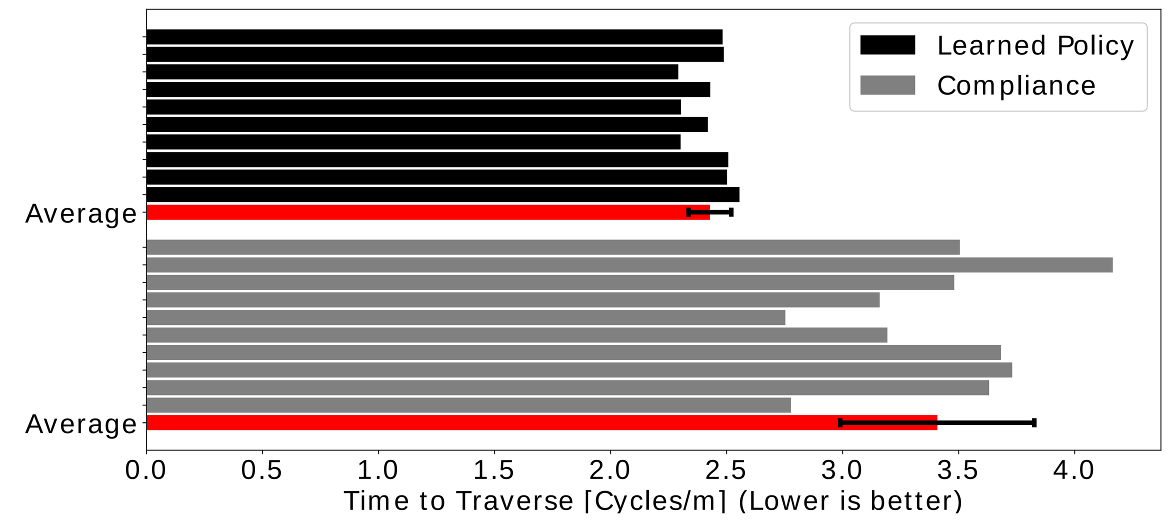

Figure 7: Learned controller outperforms the compliant controller by more than 40%

Conclusion

The proposed approach also shows that an individual agent in distributed learning framework can be assigned to an independent portion of the robot. The learned policy is tested on the robot and the performance is shown to outmatch the current state of the art by more than 40%, on a set of randomized environments.

References

[1] G. Sartoretti, Y Shi, W. Paivine, M. Travers, and H. Choset. Distributed learning for the decentralized control of articulated mobile robots. In *Submitted to 2017 International Conference on Robot Learning, CORL 2017*, 2017.

Future Research

- Apply this approach to other types of robots. Specifically, a similar approach could be used to distributively learn a policy for a walking robot.
- Investigate the possibility of applying this approach in more challenging environments such as rocks using online learning.

Contact Information

Yunfei Shi
Email yunfei.shi@connect.polyu.hk
Phone +852 9345 4060