

Spatial State Representations for Deep Reinforcement Learning, Milestone 7 15-400, Spring 2019

Joshua Zhanson
<http://jzhanson.com/15400-s19>

April 24, 2019

1 Major Changes

No major changes.

2 What You Have Accomplished Since Your Last Meeting

It seems like smaller models than the ones we have been using have a comparable training speed and performance to the larger models, which is good news. Hopefully evaluations confirm this—instead of our best performing model having 10x the parameters of our baseline, it now has a comparable number of trainable parameters.

We also have been tuning more complex bodies and writing scripts to randomly generate complex bodies. In particular, we have been testing whether a 3-joint raptor-style leg can learn to walk and we have recently found (on a simple body consisting of just a square and two raptor-style legs) very good results. The design of the leg allows the body to run up steps and slopes very naturally, jump over obstacles and pits, and even leap off of the top of obstacles to run faster with less effort within two days of training on a desktop computer. We have also completed a script to generate dog bodies, which consist of a front and a back leg (instead of two legs side-by-side) attached by flexible body segments and with a flexible neck (and head) and tail.

3 Meeting Your Milestone

I did not really meet my milestone this week, since a great deal of time was required to tune complex bodies, which consisted of a lot of trial-and-error where we would experiment with the joint angle tolerances and the motor strengths to make sure that the body could plausibly walk rather than training and going for results. I unfortunately was not able to further investigate the role of spatial and temporal information in locomotion learning.

I did, however, complete the first draft of my poster, register for ICLR, and book travel. I also emailed the department to ask about travel funding for undergraduates and am waiting on a response.

4 Surprises

As mentioned before, it was quite surprising the simple body with the 3-joint raptor leg performed so well on our “Hardcore” environment. We will have to cook up harder challenges for policies in the future.

5 Looking Ahead

At this point, we are first working on a difficult multi-body learning task involving all different kinds of bodies and then on trying to solve that task, with or without spatial information—the problem of a general multi-body locomotion policy remains, despite us shifting our focus somewhat away from spatial and temporal information. As mentioned in the previous meeting, we want to move into more structured learning (such as expert distillation) instead of blindly throwing optimization at the problem and hoping for the best, which started feeling like a lot of work for low reward.

The stellar performance of the 3-joint leg on our hardest (“Hardcore”) terrain also means that we will have to build more challenging terrains. Some possibilities are very closely spaced gaps to test counting strides and step pacing, repeated wide gaps that require leaping over, stairs of uniform or different heights and widths, taller blocks to require stopping and jumping, or steep slope ascents or descents.

Beyond those, I plan to continue work on building walker bodies and more complex randomization scripts that allow extensive mixing and matching of robot body parts. The raptor body is almost workable besides the fact that the long head, neck, and tail collapse immediately and hit the ground, ending the episode prematurely.

I also have been thinking about building a simple pygame GUI editor for robot bodies that allows a user to use the mouse and keyboard to construct a robot body, export it to JSON, and then train models on it for a short amount of time. While the implementation effort would be significant, I believe that this would be a good contribution to the community and advance the feasibility of the Box2D physics simulator as well as the usability of this walking task in reinforcement learning. I plan to start it sometime in May and work on it concurrently with the work we will do for the CoRL paper.

6 Revisions to Your Future Milestones

For Meeting of the Minds and the SPiRL ICLR workshop, the only revision we plan to make to the paper is to replace the results of our large spatial model with those of our smaller spatial model so the contrast between a baseline with 1/10th of the parameters of our proposed model is not so stark. Hopefully evaluations will turn out favorably.

I also realize that the earlier milestones I set were predicated greatly on having working bodies to train on, but that has generally not been the case. The raptor body had the issue that the leg motors initially were not powerful enough to prevent the body from collapsing and still has the issue that the head hits the ground very quickly, which needs to be remedied. I also wrote the milestones with too much breadth and not enough time to examine each body in depth or even to address the larger, more important questions of spatial and temporal information as well as a general locomotion policy. I think that these recent revision to milestones has been indicative of a shifting of focus away from the specific applicability of spatial and temporal information to attempting to construct a set of bodies upon which a general locomotion policy can emerge.

7 Resources Needed

I am currently waiting on word back from the department regarding travel funding, but I have all the resources I need.