Spatial State Representations for Deep Reinforcement Learning, Milestone 3
15-400, Spring 2019

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

February 13, 2019

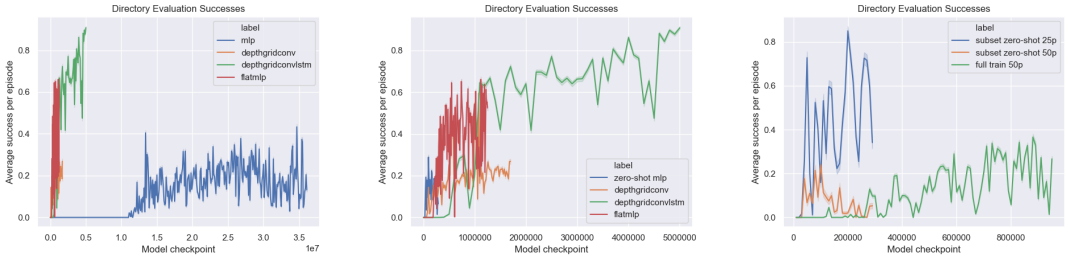

Figure 1: **Left:** Training curves for 4 models. Note the extent of baseline MLP model in comparison to the others—this is likely due to poor hyperparameters, and a better model is being evaluated. **Center:** Same graph as left, but with MLP baseline replaced with zero-shot MLP baseline trained on canonical body. **Right:** baseline MLP model trained and evaluated on full asymmetrical legs 50% dataset, with zero-shot results in 25% and 50% subsets of datasets shown for comparison.

1 Major Changes

No major changes. Pushing hard to submit to ICLR 2019 SPiRL workshop on March 7th.

2 What You Have Accomplished Since Your Last Meeting

I added the ability for the script we use to generate random bipedal walker bodies to generate bodies with asymmetrical legs and then proceeded to generate two new datasets of 30 bodies each (split into training, validation, and testing subsets) with the length of each leg segment varied up to 25% and 50%, respectively. Evaluations showed that the canonical walker policy was able to easily transfer to the unseen 25% perturbation bodies, but had somewhat more difficulty on the 50% perturbation bodies. I then generated a larger dataset of 120 50% perturbed bodies and added the original 30 bodies to this dataset and trained models for a short period of time on this dataset. Evaluations are currently in progress.

Additionally, I began writing up the results we do have, as well as details of a new positional MLP baseline model that seems to be able to achieve good performance on the original off-center hull dataset.

3 Meeting Your Milestone

As I was somewhat vague in my last milestone report about what my revised milestones for the extended abstract submission are, it is difficult to quantify to what extent I did or did not meet my milestone. However, my gut feeling is that I probably achieved 75% of what I needed for this milestone in the past two weeks. Ideally, I would have completed the discussion of the results we have today and have evaluations completed for the asymmetric leg bodies as well as being in the middle of working on the new heavy randomization environment.

4 Surprises

I did more thinking about whether training on a small, handcrafted dataset to more easily show the learning capability of our models was a good idea, and after a while (it has been about a month since I first conceived of the idea) I have decided that such training is probably not worthwhile pursuing at this point—the baseline does not have as much trouble as we expected on the full datasets, and we need to focus on tasks in which the baseline struggles in order to show improvement. Additionally, it is difficult (as I have discovered) to conceptualize a dataset that captures a balanced idea of what should be learned in as few bodies as possible—it is likely better science to randomize larger datasets. In this vein, I'm bringing up the priority of a completely random environment, where every training episode a new random body (within some constraints) is generated. Evaluations, however, become a little more hairy, since we no longer have a fixed dataset of bodies.

5 Looking Ahead

There are almost exactly three weeks until the submission deadline, which is a lot less time than I would like—but that's just the name of the game. Over the next two weeks, I hope to complete the remainder of the work necessary for the extended abstract so that the last week is spent polishing and potentially waiting on a last batch of evaluations. This entails running the models we have tentatively trained and evaluated for more time in order to have longer training curves and running evaluations on these models, finishing writing up the paper draft and then shortening (or rewriting i.e. `https://www.cs.indiana.edu/icfp96/advice.html` or `http://www.cs.sfu.ca/CourseCentral/891/abulatov/materials/How_to_write_an_extended_abstract_so_that_it_is_more_likely_to_be_accepted_POPL.p91-wegman.pdfa`) the paper for an extended abstract. Finally, there is not a lot of time left for running models, so the datasets we decide to pursue in the next week will probably be the last ones that we will be able to put in the submission. Hopefully we can tackle fully random bipedal walker bodies, randomized raptor bodies, or at least segmented bodies with off-center legs and asymmetrical legs.

6 Revisions to Your Future Milestones

No revisions besides those mentioned above. Hopefully after the paper submission there will be a few days to think about where to take the project next.

7 Resources Needed

I have all of the resources I need.