

Daily Log

Monday November 18

After running into issues trying to dual-boot on my laptop (it wouldn't let me boot from USB), I changed my mind to work on my project on the syslab computers, because I dual-booted linux on my desktop at home. It's also easier to use zoidberg on there.

Wednesday November 20

I read more into the DIAYN algorithm. It uses soft actor critic where the option is the actor and the discriminator(which tries to identify the option based on the state space it is in) is the critic. I learned about soft actor critic and found that the version in the paper it was released in didn't support discrete environments. This is a problem because the number of options at each level is finite. I read what changes were necessary. Soft actor critic uses a lot of networks. There are 2 Q functions, a value function, a target value function and the policy itself. Since soft actor critic tries to combine the strengths of q learning and policy gradients which is why many networks are needed to create more stable updates to the policy.

Friday November 22

I started to create a soft actor critic implementation. This will help me both compare the base DIAYN algorithm and my modification. By the end of class, I was able to finish most of it. By next week, I think I can finish it and create a discrete version. I forgot to push my progress on Friday. I did it as I was creating the journal.

Timeline

Date	Goal	Met
Today minus 1 week	Figure out when to make a new layer in the hierarchy	Yea because I revised my method of expansion so it does both now.
Today minus 1 week Today	Research more on state abstraction try to make HAC to run on my laptop (install dependencies and dual-boot linux on my new laptop)	This was a vague goal, but I did make progress in this area and found a solution to a problem I had. No I'm going to do it later on the sys-lab computers
Today plus 1 week	Implement soft actor critic and its discrete version. Test it on cart-pole and some discrete environment	
Today plus 2 weeks	Implement DIAYN	
Winter goal	Implement the algorithms I am comparing my model to	

Reflection

I think I am on track to finishing my winter goal. I expect that I will be able to have DIAYN and HAC working on three atari ram environments. I don't think I can implement another model by winter break, but maybe I can afterwards. While it would be nice if my algorithm works on the first try, if it doesn't work, I can't give up on it immediately because it could have failed because some hyper-parameters were making it not work. This is one of the reasons I wanted to do this project. HRL is very finicky with hyper-parameters, so the less there are, the better.