

# Appendix

## A EXPERIMENTAL DETAILS

We use Pytorch for our framework. The latent dimension is 12 in total. We use the global latent space to be two Dirichlet distribution, each of dimension 3 to model the mixture of tasks. For local latent space, we use two categorical distribution, each of dimension 3 to model a single on-going sub-task. For fair comparison, the PEARL baseline also has 12 dimension of Gaussian random variables. We train our model for 300 epochs for all the continuous control environment. Each epoch contains 4000 iterations of SAC update. For the point-robot navigation experiment, the latent space we used are 6 dimension of Gaussian distribution versus 3 categorical distribution, each of 2 dimension, versus 3 Dirichlet distribution, each of 2 dimension. Hence, the number of random variables stay the same for fair comparison.

For the continuous control tasks, the number of training and test tasks on HalfCheetah is 50 and 10; the number of training and test tasks on Humanoid is 15 and 5. Given the small set of training tasks, our task is considered harder than the previous benchmarks, where they do random sampling at the beginning of each episode. The maximum episode length are both 500. We will release our code upon publication. All the experimental results are averaged across 3 random seeds.