PhD

Decoupled Neural Interfaces using Synthetic Gradients ax1608.05343

2018/11/11 1:47 PM

- Main idea is that synthetic gradients can be used instead of real back propagated gradients to eliminate update locking so that each layer can be updated as soon as its forward pass is complete
- not much information seems to have been given about the nature of the synthetic gradients models; for the most part. these seem to be ordinary shallow networks somewhat similar to the original network that they are modelling;
- the target for these models are the synthetic gradients from the next layer which would change in the course of training though the paper itself says that these targets are fixed which is not quite clear;
- this method in general seems to be very similar to several methods from reinforcement learning including TD lambda and actor critic networks and apparently another paper has shown that many of these other methods can be shown to be special cases of synthetic gradients framework given in this paper;
- the paper only provides the results for pure synthetic gradients and pure back propagation though it also mentions some theory about a mixture of the two called BP lambda which is equivalent TD lambda where a weighted average of future gradients is incorporated with the weights decreasing exponentially similar to the exponentially discounted rewards used to compute action/state value in reinforcement learning;
- The concept of using the model network to estimate the loss itself and instead of its gradients is also mentioned though it doesn't seem to have been experimented with

related papers

2018/11/16 1:10 PM

Each layer in the network sort of behaves similarly to a state in an RL system and the corresponding synthetic gradients are equivalent to estimated state values;