Continual Model-Based Reinforcement Learning with Hypernetworks

Model-baed Reinforcement Learning (MBRL) requires learning of a dynamics model which can be computationally expensive in the case of continual learning tasks. Models need to be accurate and generalizable to scenarios consistinf of similar dynamics. To this end, the work proposes a novel scheme for training MBRL using hypernetworks for continual learning. The dynamics model is optimized using task-conditional hypernetworks which infer a task embedding from the environment and output parameters of the model. Utiliation of hypernetworks in the planning process allows task-aware dynamics and fixed capacity of networks. Furthermore, the scheme enables constant time dynamics. The proposed approach, named HyperCRL, depicts improved performance on continual MBRL tasks in comparison to meta-learning and conventional finetuning methods.

Continual learning requires updating parameters of dynamics model while tackling the problem of catastrophic forgetting. An accurate model leads to effective planning even in the presence of short horizon steps. HyperCRL aims to bridge the gap between the computational expense and accuracy in producing an accurate model for continual learning. The scheme makes use of hypernetworks which are conditioned on a task embedding obtained from the environment. Hypernetworks are trained to yield parameters of the dynamics model which is further used in planning. The planning process is carried out by making use of conventional CEM wherein the first action corresponding to the best action sequence is executed. Training on a single task is followed by the next task. However, each task consists of its own repla buffer and the robot is limited to finite mount of memory from its past experiences. In order to alleviate catastrophic forgetting, hypernetworks are regularized corresponding to all previous task embeddings.

HyperCRL is tested on two robot control scenarios consisting of pushing a block of non-uniform mass and opening a door. The scheme depicts improved performance in comparison to finetuning methods and is found to competitive to meta-learning. Suitability of hypernetworks is validated by combining them with the multi-task setup which results in improved average returns. However, the work presents several caveats. Firstly, the baselines considered in experiments do not consist of MBRL methods which would exploit the learned dynamics of the model. Secondly, the significance of hypernetworks is undermined when the body of the robot is changed to HalfCheetah resulting in sub-optimal performance. This indicates that HyperCRL may not be generalizable to all robot types. Lastly, the training procedure does not throw light on how are the task-embeddings extracted. In case they are one of agent's component then their time complexity should be taken into consideration when training hypernetworks.

The work presents two main directions for future research. (1) HyperCRL could be compared to competitive MBRL methods such as Dreamer and MBPO which train large model classes for accurate planning. (2) Extending hypernetworks to world models could be useful since a complete model of the world requires ample past experience.

Constructing accurate dynamic models which are robust to catastrophic forgetting is a challenging scenario in continual learning. The work introduces HyperCRL, an MBRL training scheme for continual learning which uses hypernetworks to construct a model during lifetime of the agent. Hypernetworks are conditioned on task embeddings and yield parameters of model which learns dynamics during different tasks executed in continuation. With access to limited past experience and tasks prone to catastrophic forgetting, HyperCRL demonstrates improved performance on block-pushing and door-opening tasks in comparison to finetuning and meta-learning methods.