Paper Critique

Anirud N, CE21B014

Course: DA7400, Fall 2024, IITM
Paper: Unveiling options with Neural Decomposition
Date: 4 September 2024

Make sure your critique addresses the following points:

1. The problem the paper is trying to address
2. Key contributions of the paper
3. Proposed algorithm/framework
4. How the proposed algorithm addressed the described problem

Note: Be concise with your explanations. Unnecessary verbosity will be penalized. Please don't exceed 2 pages.

---

1 Problem Statement

This paper addresses the existing issue in Reinforcement Learning, where the agent learns policies for a few specific tasks but doesn't generalize this knowledge to other related tasks, i.e., limiting the agent's ability to learn from the past. This lack of transferring ability in knowledge leads to lesser efficiency. This paper explores the ability of agents to transfer knowledge from one task to another by extracting "helpful" temporally extended actions(options) from neural networks. Given a neural network encoding of a task, the paper talks about decomposing this network into sub-networks that are used to create options.

2 Key Contributions

This paper provides a novel way to enhance the ability to transfer knowledge and accelerate learning from one task to another by decomposing neural networks in reusable sub-policies.

- Maps network into structures similar to oblique decision trees. Neural networks with piecewise activation functions, such as ReLU, can be represented as neural trees, where each node corresponds to a linear function of the input and each sub-node is a sub-policy.
- Wraps this sub-policy in while loops to learn options
- Uses Levin loss as a metric for selecting useful options. Lewin loss approximates an agent's policy in the early stages of learning.
- This approach is being validated by a 2-grid world problem, where the author claims that the explorations can be challenging
- The proposed method automatically learns all options components from the data generated by the agent's interaction with the environment.

3 Proposed Algorithm/Framework

- Using policy gradient approaches to learn the set of policies for the tasks
- Decompose each neural network into a set of sub-policies. The leaf nodes give the neural tree's output layer.
- The node value gives the function that each node has to follow to reach the given point. A piecewise linear activation function is used to segregate the policies into sub-policies.
- The paper defines the function of the output layer as the linear function of input X. Since every neuron represents a linear function, its combination is also a linear function
- These sub-policies now need to sample trajectories and obtain temporal abstraction. For this, we use a loop to wrap each sub-policy into the loop for some steps = z.
- Not all these options are needed to transfer knowledge. To determine only the useful options and disregard others, we use a LEVIN LOSS objective function, which estimates the number of steps needed to reach the target.
- We need to use dynamic programming to find the loss function values
- Finally, use the learned options as part of the action space to learn policies for tasks

4 Conclusions and Results

The paper shows that the options can occur "naturally" in the neural network encoding policies. Neural decompositions can be used to extract such options. The key assumption is that the decomposition approach assumes a set of related tasks where the goal is to transfer knowledge from one set of tasks to another. This means that the number of options grows exponentially, so a greedy approach is used to find the subset that minimizes the Levin loss. The paper shows that this approach accelerates the learning for the set of tasks similar to those used to train the models,