RODE: Learning Roles To Decomponse Multi-Agent Tasks

Assigning roles to decompose large multi-agent action spaces is essential for solving complex tasks. To this end, the work presents a ROle DEcomposition (RODE) scheme which learns to select and assign roles to agents based on factored action spaces. RODE first decomposes joint action spaces into restricted role spaces which are clustered and learnt by the role selector. This leads to the formulation of role selection as a bi-level hierarchy which operates by assessing the effect of each roles's actions on the environment. RODE demonstrates improved performance on 10 out of 14 StarCraft II maps in comparison to pre-existing state-of-the-art methods.

Intuitively, roles are sub-tasks which are solved by each agent corresponding to their individual policies. Prior to learning roles, joint action space is decomposed into restricted role space using clustering. The forward model learns action representations using the action encoder and predicts induced rewards and local observations to quantify the effect of actions. Following decompositions, the bi-level learning framework consists of role selection and assignment at the top level. Role selector makes use of Q-learning to learn role representations which are constructed from action representations and yield Q-values as dot products with encoded action-observation history. The lower level consists of policy exploration and execution in the assigned restricted role space. Following role assignment, an agent follows the role for fixed timesteps using its policy. Agents utilize Q-learning to learn encoded action-observation history in conjunction with action representations. The framework of Q-learning is modified to use $mixing$-networks adopted from QMIX.

RODE demonstrates state-of-the-art performance on 10 out of 14 StarCraft II micromanagement scenarios consisting of hard and super-hard maps. Suitability of role selection and assignment are validated from the improved exploration and convergence of RODE in case of tasks where conventional QMIX and ROMA present slower convergence and sub-optimal policies. Additionally, RODE is found to be suitable for transferring knowledge to maps consisting of larger number of opponents. While experiments depict improved performance and suitability, the method presents a few caveats. RODE fails to depict performance gains and sample-efficiency on easier maps which is justified by the low requirement of exploration. Unlike easy maps, RODE does not guarantee performance gains on hard maps wherein the number of agents are high and exploration is significant(such as bane_vs_bane). Lastly, knowledge transfer of RODE is validated on increasing number of agents rather than increasing difficulty of map which might aid in better evaluation.

RODE demonstrates suitability of hierarchical learning in multi-agent settings and provides new directions for future research. Exploring sample-efficient convergence of RODE on hard and super-hard maps with varying requirements of exploration is an interesting direction for future work. This would provide a better understanding of the method's shortcomings. Another suitable direction is to evaluate knowledge transfer of RODE on varying settings consisting of different enemy units, difficulty levels and structured tasks consisting of navigation and assault.

Role selection allows the multi-agent framework to tackle complex tasks by exploring relevent actions in the compact role space. The work presents RODE, a multi-agent learning algorithm which assigns and selects roles based on the effects of actions on the environment. RODE constructs a bi-level learning framework wherein the top level makes use of a role selector to learn roles based on action and role representations. The lower level learns to explore policies in the role space. RODE demonstrates state-of-the-art performance on 10 out of 14 StarCraft II scenarios and extends the scope of role selection as a hierarchy in multi-agent learning.