Contracting Actions in Self-Interested Multi-Agent Reinforcement Learning

Paper #...

ABSTRACT

Whereas agents in cooperative environments share the same reward function in self-interested settings agents maximize their individual returns. As a consequence agents learned policies tend to be overly greedy especially when shared resources between agents are scarce. Cooperative game theory provides a theoretical framework for self-interested settings where binding agreements between agents are possible. In this work we propose binding agreements for multi-agent reinforcement learning which allow agents to transfer reward in exchange for following a non greedy trajectory. Thus, agents are enabled to compensate each other for their occurred losses that arise from behaving non-greedy. We evaluate our proposed method in a smart-factory scenario where agents compete for available machines. The tasks either have low-priority or high-priority, giving small rewards or high-rewards respectively. We empirically demonstrate that RL agents stably learn agreements and thus originate higher returns compared with the pure self-interested case.

KEYWORDS

multi-agent learning; cooperative game theory

1 INTRODUCTION

2 BACKGROUND

3 RELATED WORK

4 CONTRACTING

Algorithm 1 Contracting

procedure $\textsc{Contract}(Q^{\text{GREEDY}}, Q^{\text{NON-GREEDY}}, N_c)$
  Observe current state s_1
  Calculate compensation $q_c \leftarrow max_a(Q^{non\text{-}greedy}(s_1, a))$
  for $t = 1, N_c$ **do**
    $a_{greedy} \leftarrow argmax_a(Q^{greedy}(s, a))$
    $a_{non\text{-}greedy} \leftarrow argmin_a(Q^{non\text{-}greedy}(s, a))$
    Execute $a_t \leftarrow \langle a_{greedy}, a_{non\text{-}greedy} \rangle$
    Observe rewards $r_{g,t}, r_{n,t}$ and new state s_{t+1}
    $s \leftarrow s_{t+1}$
  Final compensation $q_c \leftarrow q_c - max_a(Q^{non\text{-}greedy}(s, a))$
  Final returns $R_g \leftarrow -q_c + \sum_{t=1}^{N_c} r_{g,t}, R_n \leftarrow q_c + \sum_{t=1}^{N_c} r_{n,t}$

5 EXPERIMENTS

5.1 Evaluation Environments

5.1.1 *Smart Factory (SF).*

5.2 Methods

5.3 Results

5.4 Discussion

6 CONCLUSION AND FUTURE WORK

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 2019, Montreal, Canada

https://doi.org/doi