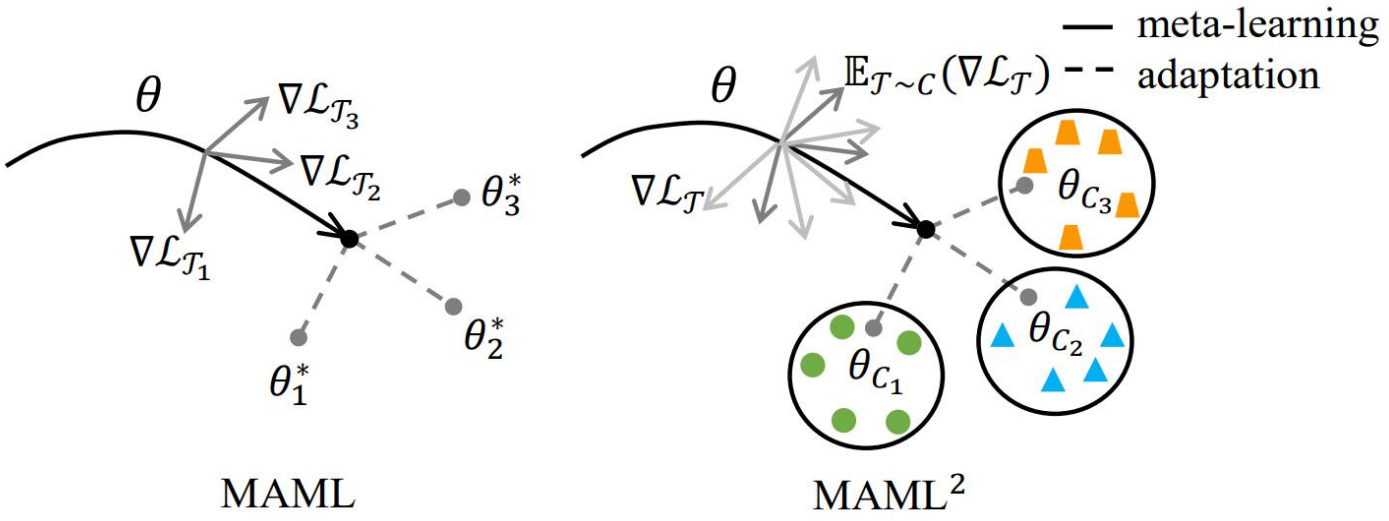
MAML^2: Meta Reinforcement Learning via Meta-Learning for Task Categories

Qiming FU, Zhechao WANG, Nengwei FANG, Bin XING, Xiao ZHANG, Jianping CHEN

Frontiers of Computer Science, DOI: 10.1007/s11704-022-2037-1

Problems & Ideas

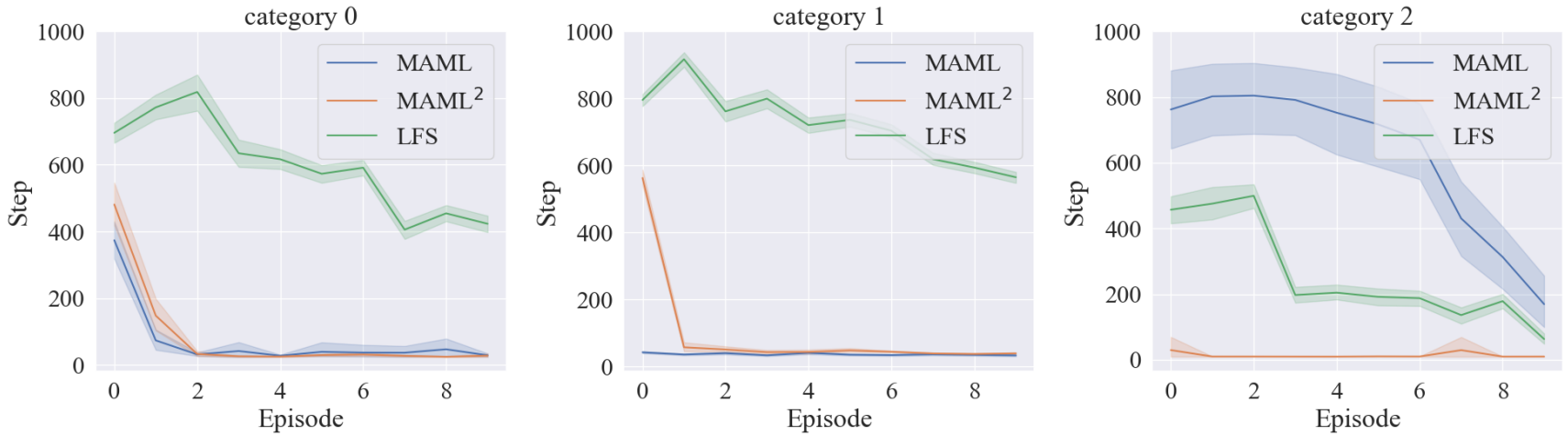
- Problems of conventional Meta-learning:
  - Ignoring some isolated tasks in pursuit of the average performance.
  - Resulting in negative adaptation in these isolated tasks.
- Ideas: A hierarchical framework of double meta-learning, and the whole framework includes classification, meta-learning, and re-adaptation.



MAML vs. MAML^2. Each circle means a task subset. MAML updates the meta-parameter with the gradients from multiple task learnings whose optimal parameters are θ^*, while MAML^2 updates the meta-parameter with the gradients from multiple task subsets learning whose category parameters are θ_C. Each θ_C refers to the meta-parameter of the corresponding task subset.

Main Contributions

- Contributions:
  - A classification method based on learned parameter similarities to separate out some isolated tasks;
  - A hierarchical framework of double meta-learning to train a more robust initial parameter;
  - A re-adaptation process to accelerate the whole learning process in a new task, which resets the temporal parameter with a category parameter by the ‘update tendency’.



Adaptation process. (1) For those un-isolated tasks (in category 0 and 1), although MAML^2 presents a weak initial performance in the first episode, a close performance like MAML will be achieved after the re-adaptation process (after episode 1); (2) For isolated tasks (in category 2), MAML^2 has an excellent initial performance, while MAML always occurs a negative adaptation in many trials. MAML^2 successfully prevents the agent from negative adaptation without losing the average performance. LFS refers to learning from scratch.