Noname manuscript No.
(will be inserted by the editor)

Survey - Asynchronous Methods for Reinforcement Learning

Johannes Czech

Received: date / Accepted: date

Abstract This paper aims to address the issue of often high computational requirements for Reinforcement Learning application by validating techniques introduced by asynchronous methods. First go over the general principle and the motivation behind asynchronous methods. Brief history when asynchronous where used for RL the first time. Next cover the theory of techniques of shared weight updates.Taking technical challenges into account such as process communication and memory requirements. Description and discussion of the super-linearity phenomena when using asynchronous methods. Discuss application areas for asynchronous methods and advantages and potential problems when using them.

Keywords Reinforcement Learning · Asynchronism · Super-linearity

1 Introduction

From a historical perspective asynchronous have appeared very early in the research. One of the earliest use cases of asynchronous methods for RL was in the context of Dynamic Programming (Bertsekas and Tsitsiklis 1989 -¿ reference). Later this technique was built upon in an algorithm called Prioritized Sweeping -¿ reference which is acts the same way as the Gauss Seidel algorithm for Dynamic Programming but concentrates all computational effort on the most "interesting" parts of the system. *-¿ shortly go over Dynamic Programming (Bellman 1957) -¿ Bellman's optimality equation*

F. Author
first address
Tel.: +123-45-678910
Fax: +123-45-678910
E-mail: fauthor@example.com

S. Author
second address

An important result from the theory of Markov decision tasks tells us that there always exists at least one policy which is *optimal* in the following sense. For every state, the expected discounted reward-to-go using an optimal policy is no worse than that from any other policy" (S.14, "Prioritized Sweeping"). Makes use of heuristic, introducing a hyperparameter -¿ Question: What is the main application of asynchronous methods here?

Moreover Dimitri P. Bertsekas extended his ideas in a "Distributed Asynchronous Policy Iteration in Dynamic Programming" 2010. -¿ reference, "Asynchronous Methods for Deep Reinforcement Learning" 2016 -¿ reference

Your text comes here. Separate text sections with Motivation History Theory Mathematical Formulation

The revival of asynchronous methods has been started by DeepMind Massively Parallel Methods for Deep Reinforcement Learning -¿ reference

2 Taxononomy of distributed reinforcement learning

Taxonomy -¿ Paper: Communication Efficient Distributed Reinforcement Learning

Approaches Q-Learning, Policy-Gradient Methods, Actor-Critic Methods

2 settings: * multi-agent collaborative RL – multiple agents aim to maximize the team-averaged long-term reward via collaboration in a common environment * parallel RL – multiple parallel machines are used for solving large-scale MDPs with larger computational power and higher data efficiency

Distributed Methods are applicable in various RL methods * Policy Gradient * Q-Learning (distributed Q-Learning) * Actor-Critic Approaches (A3C / A2C)

-¿ use a figure here to describe the concept

Problem formulation: Multi-agent collaborative RL: * each agent m observes globabl state s_t e S shared by all agents and takes action a_m,t e A_m local action is determined by local policy

Problem formulation: Parallel reinforcement learning * solving large-scale single agent RL task, running in parallel on multiple computation units e.g. workers * advantage training time reduction, training stabilization * aim to learn common policy for different instance of identical MDP -¿ losses and initial state is different for each worker

3 Superlinearity phenomena

Text with citations [2] and [1]. Try to give proof for this phenomena otherwise rely on empirical experiments

Table 1 Please write your table caption here

first	second	third
number	number	number
number	number	number

4 Applications

Describe application areas Asynchronous learning with multiple agents Swarm robots

as required. Don't forget to give each section and subsection a unique label (see Sect. 3).

Reference: Deep Reinforcement Learning for Robotic Manipulation with Asynchronous Off-Policy Updates

* Addresses problem of high sample complexity for reinforcement learning tasks * Makes use of off-policy training of deep Q-functions * Learning of complex control tasks without prior demonstration, from their knowledge the first use case of this * Speed up learning by using asynchronous methods * usage of Deep Deterministic Policy Gradient algorithm (DDPG) and Normalized Advantage Function algorithm (NAF) achieve training times suitable for real robotic systems * introduction of new asynchronous variant of NAF * focuses on model-free RL * "The goals of this prior work are fundamentally different from ours: while prior asynchronous deep reinforcement learning work seeks to reduce overall training time, under the as- sumption that simulation time is inexpensive and the training is dominated by neural network computations, our work instead seeks to minimize the training time when training on real physical robots, where experience is expensive and computing neural network backward passes is comparatively cheap." * don't make use of replay buffer but asynchronous updates instead collecting experience across multiple robotic platforms * achieved significant speedup in overall training time

Collective Robot Reinforcement Learning with Distributed Asynchronous Guided Policy Search * making use of asynchronous updates to have greater quantity and diversity of experience * task visual-based door opening task * guided policy search simultaneously trained with a repay buffer of the latest trajectory samples

Paragraph headings Use paragraph headings as needed.

$$a^2 + b^2 = c^2 \tag{1}$$

5 Conclusion

Advantages when using async updates Disadvantage problems, limits for using asynchronous methods

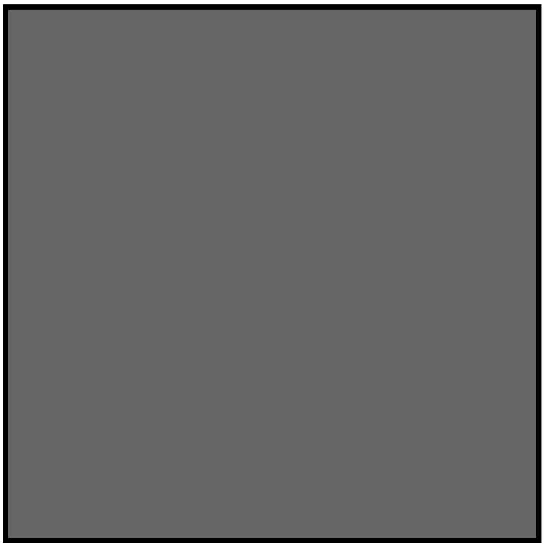

Fig. 1 Please write your figure caption here

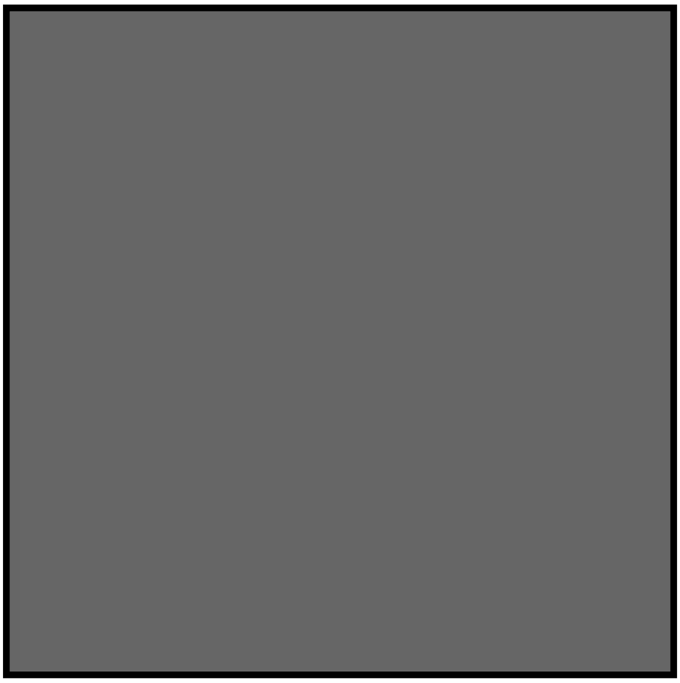

Fig. 2 Please write your figure caption here

References

1. Author, Article title, Journal, Volume, page numbers (year)
2. Author, Book title, page numbers. Publisher, place (year)