

Reinforcement Learning

Current Trends and Future Directions

Book of Abstracts – Day 3 (Feb 28)

14. Prashanth L. A., IITM, Chennai

Title: A Cubic-regularized Policy Newton Algorithm for Reinforcement Learning

Timings: 28 Feb, 9 – 10 am

Abstract: We consider the problem of control in the setting of reinforcement learning (RL), where model information is not available. Policy gradient algorithms are a popular solution approach for this problem and are usually shown to converge to a stationary point of the value function. In this paper, we propose two policy Newton algorithms that incorporate cubic regularization. Both algorithms employ the likelihood ratio method to form estimates of the gradient and Hessian of the value function using sample trajectories. The first algorithm requires an exact solution of the cubic regularized problem in each iteration, while the second algorithm employs an efficient gradient descent-based approximation to the cubic regularized problem. We establish convergence of our proposed algorithms to a second-order stationary point (SOSP) of the value function, which results in the avoidance of traps in the form of saddle points. In particular, the sample complexity of our algorithms to find an ϵ -SOSP is $O(\epsilon^{-3.5})$, which is an improvement over the state-of-the-art sample complexity of $O(\epsilon^{-4.5})$.

15. Sridhar Mahadevan, UMass, Adobe, USA

Title: Universal Imitation Games: The (Co) End of Generative AI

Timings: 28 Feb, 10 – 11 am

Abstract: In this final talk, we go over two profound ideas by the mathematician Yoneda. In 1954, the year of Turing's death, the Yoneda Lemma was discovered in a Paris train station. It simply states that objects in a category can be defined from their interactions. We apply the Yoneda Lemma to construct universal representers of objects in non-symmetric generalized metric spaces. We discuss the connection of Yoneda

embeddings to define more general attention models in LLMs. We then introduce Yoneda's categorical integral calculus, and show how to construct the geometric Transformer model by computing the coend of the nerve of the category of Transformer models.

16. Prateek Jain, Google Research, Bengaluru

Talk Details: TBA

Timings: 28 Feb, 11:30 – 12:30 pm

17. Jayakumar S, Adobe, Bengaluru

Title: Sequential decision making in large systems – RL with partial observability, multiple agents and dynamic environments

Timings: 28 Feb, 12:30 – 1:30 pm

Abstract: Reinforcement learning in domains such as marketing offers several challenges, the key ones being partial observability, presence of multiple agents, very high-dimensional action spaces and the dynamic nature of the environment itself. In this talk we look at different sub-problems and possible solutions that could be integrated to solve the main problem. These sub-problems can be broadly divided into two categories – learning to simulate the environment and learning to act optimally in the simulated environment. For both these categories we will look at some theoretical guarantees and some practical approaches using diffusion models and transformers in general and large language models in particular and will finally discuss putting all these together into a decision making solution.