

**CRUNCH Seminars at Brown, Division of Applied Mathematics**

**Friday - April 16, 2021**

**Distributional Offline Continuous-Time Reinforcement  
Learning with Neural Physics-Informed PDEs**

**By Igor Halperin, AI AM Research, Fidelity Investments**

**This paper addresses distributional offline continuous-time reinforcement learning (DOCTR-L) with stochastic policies for high-dimensional optimal control. A soft distributional version of the classical Hamilton-Jacobi-Bellman (HJB) equation is given by a semilinear partial differential equation (PDE). This 'soft HJB equation' can be learned from offline data without assuming that the latter correspond to a previous optimal or near-optimal policy. A data-driven solution of the soft HJB equation uses methods of Neural PDEs and Physics-Informed Neural Networks developed in the field of Scientific Machine Learning (SciML). The suggested approach, dubbed 'SciPhy RL', thus reduces DOCTR-L to solving neural PDEs from data. Our algorithm called Deep DOCTR-L converts offline high-dimensional data into an optimal policy in one step by reducing it to supervised learning, instead of relying on value iteration or policy iteration methods. The method enables a computable approach to the quality control of obtained policies in terms of both their expected returns and uncertainties about their values.**