

CREST - GENES

Cours doctoraux 2025 – 2026

INTRODUCTION TO CONTINUOUS-TIME REINFORCEMENT LEARNING

XUNYU ZHOU

Columbia University

Industrial Engineering & Operations Research Department

SCHEDULE	Monday	9 March 2026 16 March 2026	1:30 PM – 5:30 PM	ROOM 2006
	Thursday	12 March 2026	1:30 PM – 5:30 PM	ROOM 2006

Content

This crash course covers fundamental theory and algorithms for reinforcement learning with continuous-time controlled diffusion processes, which have been developed in the last five years. It includes the following topics.

1. Exploration vs exploitation: relaxed control, entropy regularization, exploratory HJB equation and Gibbs measure.
2. Gaussian exploration under linear-quadratic control: optimality of Gaussian exploration and cost of exploration.
3. Temperature control of Langevin diffusions: simulated annealing for nonconvex optimization and optimal temperature control.
4. Policy evaluation: martingale characterization, martingale loss function and martingale orthogonality conditions.
5. Policy gradient: policy gradient via policy evaluation, and temporal difference learning. q-learning theory: generalized Hamiltonian and policy improvement, Q-function and q-function, martingale characterization of q-function.

Evaluation

Take home project.

References