Mini Course | קורס מרכז

**REINFORCEMENT LEARNING: INTRODUCTION TO A COMPUTATIONAL FRAMEWORK FOR UNDERSTANDING DECISION MAKING**

Animals (including humans) base their decisions on past learning from trial and error. Bringing ideas from computer science to bear on neuroscience and psychology, since the 1990’s reinforcement learning theory has provided a precise, interpretable computational framework for understanding both learning and decision making. In this mini course, I will give an introduction, from the ground up, to this framework and how it can be used to ask and answer questions about behavior and the brain. I will start with Pavlovian (classical) conditioning, its modeling as prediction learning, and its implementation in dopamine-dependent learning in the basal ganglia. Then we will discuss instrumental (operant) conditioning and different algorithms for adaptive action selection in reinforcement learning, including model-based and model-free learning. In the third and final part of the mini-course, we will discuss extensions based on interests of course attendees. Potential topics include modeling response rates, applications to mental illness, and representation learning.

---

**Abstract**

Animals (including humans) base their decisions on past learning from trial and error. Bringing ideas from computer science to bear on neuroscience and psychology, since the 1990’s reinforcement learning theory has provided a precise, interpretable computational framework for understanding both learning and decision making. In this mini course, I will give an introduction, from the ground up, to this framework and how it can be used to ask and answer questions about behavior and the brain. I will start with Pavlovian (classical) conditioning, its modeling as prediction learning, and its implementation in dopamine-dependent learning in the basal ganglia. Then we will discuss instrumental (operant) conditioning and different algorithms for adaptive action selection in reinforcement learning, including model-based and model-free learning. In the third and final part of the mini-course, we will discuss extensions based on interests of course attendees. Potential topics include modeling response rates, applications to mental illness, and representation learning.