
NORM FERNS, York University

Bisimulation Through Markov Decision Process Coupling

Markov decision processes (MDPs) are a popular mathematical model for sequential decision-making under uncertainty. Many standard solution methods are based on computing or learning the optimal value function, which reflects the expected return one can achieve in each state by choosing actions according to an optimal policy. In order to deal with large state spaces, one often turns to approximation.

Bisimulation metrics have been used to establish approximation bounds for state aggregation and other forms of value function approximation in MDPs. In this talk, we show that a bisimulation metric defined on the state space of an MDP in previous work can be viewed as the optimal value function of an optimal coupling of two copies of the original model, and discuss the consequences thereof.

This is joint work with Doina Precup.