Title:
Shielded Reinforcement Learning for Hybrid Systems

Abstract:
In this talk, I present recent work (currently under review) with colleagues in Aalborg.

Safe and optimal controller synthesis for switched-controlled hybrid systems, which combine differential equations and discrete changes of the system's state, is known to be intricately hard. Reinforcement learning has been leveraged to construct near-optimal controllers, but their behavior is not guaranteed to be safe, even when it is encouraged by reward engineering. One way of imposing safety to a learned controller is to use a shield, which is correct by design. However, obtaining a shield for non-linear and hybrid environments is itself intractable. We propose the construction of a shield using the so-called barbaric method, where an approximate finite representation of an underlying partition-based two-player safety game is extracted via systematically picked samples of the true transition function. While hard safety guarantees are out of reach, we experimentally demonstrate strong statistical safety guarantees with a prototype implementation and Uppaal Stratego. Furthermore, we study the impact of the synthesized shield when applied as either a pre-shield (applied before learning a controller) or a post-shield (only applied after learning a controller). We experimentally demonstrate superiority of the pre-shielding approach. We apply our technique on a range of case studies, including two industrial examples, and further study post-optimization of the post-shielding approach.