



Bradley Hayes University of Colorado

“Achieving Safe Human-Robot Collaboration through Explainable AI”



*Wednesday, November 8, 2017
Noon-1:00
CIT 477 Lubrano Conference Room*

Abstract: Robots capable of fluent collaboration with humans will bring transformative changes to the way we live and work. In domains ranging from healthcare to domestic tasks to manufacturing, particularly under conditions where modern automation is ineffective or inapplicable, human-robot teaming can be leveraged to increase efficiency, capability, and safety. Despite this, the deployment of collaborative robots into human-dominated environments remains largely infeasible due to the myriad challenges involved in creating helpful, safe, autonomous teammates. In this talk I will present recent work toward overcoming these challenges, realizing flexible, communicative robot collaborators that both learn and dynamically assist in the completion of complex tasks through the application of novel learning and control algorithms. In particular, I will be focusing on the importance of transparent design and the human-interpretable models that underpin these methods, helping to ensure safe and efficient operation.

Brad Hayes is an Assistant Professor of Computer Science at the University of Colorado Boulder. He received his PhD from Yale University and completed a postdoctoral position at MIT where his work focused on enabling close-proximity human-robot collaboration, creating algorithms to enable robotic teammates to learn from and work productively with humans by anticipating their needs. Brad's research centers around developing the algorithms and explainable Artificial Intelligence techniques necessary for creating supportive, interactive, communicative, and intuitive robotic systems that are capable of establishing shared expectations with people and safely performing complex collaborative tasks in human-populated environments.

Host: Maartje de Graaf/HCRI