

Research Presentation

Minwoo Lee

Colorado State University

Monday, March 13, 2017 12:45 – 1:35 p.m., McGraw 122

Brief Q&A to Follow

Title: Sparse Bayesian Reinforcement Learning for Robust and Efficient Learning

Abstract:

An intelligent agent often fails to learn when it encounters a new environment because it forgets what it has learned. This impacts the stability of learning process. One instance of this is well-known “catastrophic forgetting” of neural networks. Biased input data used for neural network training cause the networks to lose the learned knowledge as new data arrive from outside the old data space. To overcome this, sparse Bayesian reinforcement learning proposes a way to gradually gain knowledge about the domain with a sparse learning model for efficiency. Robust learning can be achieved through a knowledge augmentation framework. With the proposed heuristics, we claim that the sparsity of the domain is maintained. The Bayesian model makes learning more efficient by providing a way to make a correct decision depending on the confidence of estimation. The proposed framework has been successful applied to analytical learning and efficient search for transfer learning and efficient fine control. This will be demonstrated in application to control of a simulated octopus arm.

Bio:

Minwoo Lee is a PhD candidate in computer science at Colorado State University. He received his M.S. and B.S. from Korea Aviation University in 2002. His research interests focus on machine learning, with emphasis on reinforcement learning, transfer learning, sparse learning, multiagent learning, fine control based policy development, knowledge representation and robust knowledge augmentation. He received REUSSI/NSF funds for research with Inria, France in 2012 and investigated robust high dimensional clustering models for biomedical data. His recent research about pretraining deep networks for reinforcement learning agents received the best paper award from IJCNN 2015. He has served as a referee for journals such as IEEE Transactions on Neural Networks and Learning Systems (TNNLS) and conferences such as IEEE Conference of Decision and Control (CDC), IEEE International Joint Conference on Neural Networks (IJCNN), and IEEE International Conference on Tools with Artificial Intelligence (ICTAI).