



**Department of Electronic Engineering  
The Chinese University of Hong Kong**

**SEMINAR**

**Sample-Efficient and Stable Deep Reinforcement Learning for Robotics**

By  
**Mr. Shixiang Gu**

**Department of Engineering  
University of Cambridge, UK**

**Date :** 16<sup>th</sup> Jan, 2017 (Monday)  
**Time :** 10:30am – 11:30 am  
**Place :** LT 1, Science Centre, CUHK

**Abstract:**

Model-free deep reinforcement learning (RL) methods have been successful in a wide variety of simulated domains. However, a major obstacle facing deep RL in the real world is the high sample complexity of such methods. We present two independent lines of work to address this fundamental problem. In the first part, we explore how off-policy deep RL methods based on normalized advantage functions (NAF) can learn real-world robotic manipulation skills, with multiple robots simultaneously pooling their experiences. Our results show that we can obtain faster training and, in some cases, converge to a better solution when training on multiple robots, and we show that we can learn a real-world door opening skill with deep neural network policies using about 2.5 hours of total training time with two robots. In the second part, we present Q-Prop, a novel model-free method that combines the stability of unbiased policy gradients with the efficiency of off-policy RL. We analyze the connection between Q-Prop and existing model-free algorithms, and use control variate theory to derive two variants of Q-Prop with conservative and aggressive adaptation. We show that conservative Q-Prop provides substantial gains in sample efficiency over trust region policy optimization (TRPO) with generalized advantage estimation (GAE), and improves stability over deep deterministic policy gradient (DDPG), the state-of-the-art on-policy and off-policy methods, on OpenAI Gym's MuJoCo continuous control environments.

**Biography:**

Shixiang (Shane) Gu started PhD in Machine Learning under Cambridge-Tübingen PhD Fellowship in the fall 2014, where he is co-supervised by Richard E. Turner and Zoubin Ghahramani at University of Cambridge, and Bernhard Schölkopf at the Max Planck Institute for Intelligent Systems in Tübingen. He also collaborates closely with Sergey Levine at UC Berkeley/Google Brain and Timothy Lillicrap at DeepMind. He obtained his B.ASc. in Engineering Science from the University of Toronto in 2013, where he completed his thesis with Geoffrey Hinton. He is funded by NSERC and a Google Focused Research Award.

**\*\*\* All are welcome to attend \*\*\***

*For inquires, please contact Prof. X G Wang, Tel. No. 3943 8283*