THE UNIVERSITY OF CHICAGO

Computer Science Department

CERES UNSTOPPABLE SPEAKER SERIES

Wei Li
Vice President and General Manager of Machine Learning and Translation, Intel Corporation

October 18, 2:30 pm, JCL 390

"The World's Deep Learning Runs on CPUs"

Abstract:

The conventional wisdom is that deep learning (DL) has to be done on accelerators. However, the world's deep learning runs very well on general purpose CPUs today, adaptable to deep learning that continues to evolve quickly from its early success in CNNs, to the adoption of RNNs, Recommendation Systems, and Reinforcement Learning with a wide range of performance characteristics for computer systems in production environments. In this talk, I will cover the hardware and software technologies that power AI in the world's clouds and data centers. I will show a 200X performance improvement, and how to reduce inference to milliseconds and scale training to 1000s of compute nodes.

Bio:

Wei Li is Vice President and General Manager of Machine Learning and Translation at Intel Corporation. In his 20 years at Intel, he has led innovative software efforts for artificial intelligence, clouds/data centers, mobile devices, and hardware/software co-design. He served as an associate editor for ACM Transactions on Programming Languages and Systems, and was on the joint Intel/Microsoft technical committee that funded the creation of parallel computing research centers at the University of California-Berkeley and the University of Illinois at Urbana-Champaign. He received a Ph.D. in Computer Science from Cornell University, and completed an executive program at Stanford Graduate School of Business.

CERES
Center for Unstoppable Computing

The talk is at 2:30pm, Thursday, October 18, Crerar 390

Refreshments after talk, Crerar 390
Host: Andrew A. Chien
Contact: 773-702-3508
http://www.cs.uchicago.edu/events