CS Colloquium

22 September, 11:30-13:00

«Algorithm Design Automation»

Zhi-Quan Luo
The Chinese University
of Hong Kong, Shenzhen

This talk addresses the challenge of designing and optimizing algorithms under strict computational and memory constraints, with applications spanning massive MIMO systems, wireless communication, and large-scale AI training. Beginning with a finite-horizon optimization perspective, we review classical gradient descent, its limitations with constant step sizes, and optimal finite-step schemes derived from Chebyshev minimax polynomials. We then present recent advances in matrix multiplication, including AI-discovered state-of-the-art algorithms for structured products such as XX^T, achieving notable speedups and energy savings over recursive Strassen methods in both CPU and GPU settings. The discussion extends to assessing large language models' (LLMs) capabilities in mathematical reasoning and novel problem solving, highlighting cases where LLM-assisted approaches led to breakthroughs. Finally, we introduce AlphaEvolve, a code-space search framework for automated algorithm discovery, demonstrating its success in improving long-standing algorithmic bounds and generating efficient CUDA kernels. The talk concludes with potential future directions, including new algorithms for causal attention, constrained SVD, and advanced GPU kernels.

Prof. Luo's talk will also include a part where he will introduce CUHK-Shenzhen University and postgraduate study opportunities there.

Pokrovsky Boulevard 11
R406