Title: Properly learning decision trees in almost polynomial time

Speakers: Guy Blanc (Stanford) and Jane Lange (MIT)

Abstract: We consider the problem of designing query strategies for black-box functions. Any adaptive query strategy for evaluating a boolean function f can be represented as a decision tree. The goal is to minimize the depth of the tree, corresponding to the number of queries to x required to evaluate $f(x)$. We give an $n^{O(\log \log n)}$-time algorithm for finding a minimum-depth tree.

In the language of learning theory, we give an $n^{O(\log \log n)}$-time membership query algorithm for properly learning decision trees under the uniform distribution over $\{\pm 1\}^n$. Prior to our work, the previous fastest runtime was $n^{O(\log n)}$, a consequence of a classic algorithm of Ehrenfeucht and Haussler.

To analyze our algorithm, we prove a new structural result for decision trees that strengthens a theorem of O'Donnell, Saks, Schramm, and Servedio. While the OSSS theorem says that every decision tree has an influential variable, we show how every decision tree can be "pruned" so that every variable in the resulting tree is influential.

This is joint work with Mingda Qiao and Li-Yang Tan. Paper available at: https://arxiv.org/abs/2109.00637