Abstract: In the string trace reconstruction problem, there is an unknown binary string, and we observe noisy samples of this string after it has gone through a deletion channel. This deletion channel independently deletes each bit with constant probability $q$ and concatenates the remaining bits. The goal is to learn the original string with high probability using as few traces as possible, where the sample complexity is characterized by the length of the string, $n$.

String trace reconstruction is arguably the hottest problem in information theory, and researchers are incredibly stuck at bridging the gap between the exponential upper bound and $n^2$ lower bound. In the first part of this talk, I will explain the string trace reconstruction problem and the current best upper bound techniques. In the latter part, I will discuss several generalizations and variants of trace reconstruction, including our work on tree trace reconstruction. For many classes of trees, including complete trees and spiders, we provide algorithms that reconstruct the labels using only a polynomial number of traces. This exhibits a stark contrast to known results on string trace reconstruction, which require exponentially many traces, and where a central open problem is to determine whether a polynomial number of traces suffice.

Based on joint work with Miklos Z. Racz and Cyrus Rashtchian

Monday, November 4 at 3:00 PM in 612 SEO