Sublinear Algorithms for Edit Distance

Barna Saha ✉

University of California at Berkeley, CA, USA

─── **Abstract** ───

The edit distance is a way of quantifying how similar two strings are to one another by counting the minimum number of character insertions, deletions, and substitutions required to transform one string into the other. A simple dynamic programming computes the edit distance between two strings of length n in $O(n^2)$ time, and a more sophisticated algorithm runs in time $O(n + t^2)$ where t is the distance (Landau, Myers and Schmidt, SICOMP 1998). In pursuit of obtaining faster running time, the last couple of decades have seen a flurry of research on approximating edit distance, including polylogarithmic approximation in near-linear time (Andoni, Krauthgamer and Onak, FOCS 2010), and a constant-factor approximation in subquadratic time (Chakrabarty, Das, Goldenberg, Koucký and Saks, FOCS 2018). In this talk, we will discuss recent progress that goes beyond linear time, and studies sublinear time algorithms for edit distance. We will also discuss the role preprocessing might play in designing fast algorithms.

This is a joint work with Elazar Goldenberg, Tomasz Kociumaka, Robert Krauthgamer, and Aviad Rubinstein.