Crossing the Logarithmic Barrier for Dynamic Boolean Data Structure Lower Bounds

Abstract:

We prove the first super-logarithmic lower bounds on the cell probe complexity of dynamic *boolean* (a.k.a. decision) data structure problems, a long-standing milestone in data structure lower bounds. We introduce a new technique and use it to prove a \( \sim \log^{1.5}(n) \) lower bound on the operational time of a wide range of boolean data structure problems, most notably, on the query time of dynamic range counting over \( \mathbb{F}_2 \) ([Patrascu07]). Proving a super-logarithmic lower bound for this problem was explicitly posed as one of five important open problems in the late Mihai Patrascu's obituary [Tho13]. This result also implies the first super-logarithmic lower bound for the classical 2D range counting problem, one of the most fundamental data structure problems in computational geometry and spatial databases. We derive similar lower bounds for boolean versions of dynamic polynomial evaluation and 2D "rectangle stabbing", and for the (non-boolean) problems of range selection and range median. Our technical centerpiece is a new way of "weakly" simulating dynamic data structures using efficient *one-way* communication protocols with small advantage over random guessing. This simulation involves a surprising excursion to low-degree (Chebychev) polynomials which may be of independent interest, and offers an entirely new algorithmic angle on the "cell sampling" method of Panigrahy et al. [PTW10].

Joint work with Kasper Green-Larsen and Huacheng Yu.