The Geometry of Binary Search Trees

Erik Demaine
MIT

Dion Harmon
NECSI

John Iacono
Poly Inst. of NYU

Daniel Kane
Harvard

Mihai Pătraşcu
IBM Almaden
Ass

- Point set in the plane
- \textbf{ASS} = any nontrivial rectangle spanned by two points contains another point (interior or on boundary)
MinASS Problem

Problem: Given a point set, find the minimum ASS superset

- Up to constant factors, can assume input points are in general position (no two horiz/vert aligned)
MinASS in Worst Case

- $O(n \lg n)$ points always suffice
- $\Omega(n \lg n)$ points are sometimes necessary

(random; bit-reversal permutation matrix)
NP-completeness

- **Theorem:** MinASS is NP-complete
- **OPEN:** General position MinASS

(a) Wire and split gadget.
(b) Shift gadget.
(c) Crossover gadget.
(d) Not-All-Equal clause gadget.
Approximating MinASS

- **OPEN**: $O(1)$-approximation?
- **Known**: $O(lg \ lg n)$-approximation [and it’s not easy]
GreedyASS

- Imagine points arriving row by row
- Add necessary points on the new row to remain ASS

Conjecture: $O(1)$-approximation
Binary Search Tree (BST)

- Recall our good old friends:
- Unit-cost operations:
  - Move finger up/left/right
  - Rotate left/right
The Best BST

Problem:
- Given (offline) access sequence $S = x_1, x_2, ..., x_m$
- $OPT(S)$ = minimum sequence of unit-cost ops in BST to touch $x_1, x_2, ..., x_m$ in order

Without rotations, this problem is solved by (static) optimal BSTs of Knuth [1971]

Ultimate goal:
$O(1)$-competitive online algorithm
Example of BST Execution

- Access sequence: 1, 4, 5, 7, 6, 2, 3
Example of BST Execution

- Access sequence: 1, 4, 5, 7, 6, 2, 3
ASS/BST Equivalence

- In fact, any ASS point set is a BST execution!
  - Treap by next access time

- **Corollary:** \( \text{MinASS}(S) = \text{OPT}(S) \)
Attacks on OPT(S)

- **Splay trees** [Sleator & Tarjan 1983]
  - Conjecture: $O(1)$-competitive
  - Many nice properties known, but no $o(lg n)$-competitiveness known

- **Tango trees** [Demaine, Harmon, Iacono, Pătrașcu 2004]
  - $O(lg lg n)$-competitive

- **GreedyASS** [Lucas 1988; Munro 2000]
  - Proposed as offline BST algorithm ("Order by Next Request")
  - No $o(n)$-competitiveness known
Online ASS/BST Equivalence

- **Theorem:** If ASS sequence computed online row by row (like GreedyASS), then can convert to an online BST algorithm with constant-factor slowdown
  - Transform to geometry and back
- **Main ingredient:** *split trees*
  - Support any sequence of $n$ splits in $O(n)$ time
  - Splay trees take $O(n \alpha(n))$ time [Lucas 1988]
  - Splay trees take $O(n \alpha^*(n))$ time [Pettie 20??]
New Dynamic Optimality Conjecture

- GreedyASS is now an online algorithm!
- **Conjecture:**
  GreedyASS is $O(1)$-competitive
- Previously conjectured to be an offline $O(1)$-approximation to OPT

[Lucas 1988; Munro 2000]
Lower Bounds

- Need lower bounds on $\text{OPT}(S)$ to compare algorithms (like GreedyASS) against
- Wilber [1989] proved two lower bounds:
  - Wilber I used for $O(\lg \lg n)$-competitiveness of Tango trees
  - Wilber II used for key-independent optimality [Iacono 2002]
- Conjecture: Wilber II $\geq$ Wilber I
- Conjecture: $\text{OPT}(S) = \Theta(\text{Wilber II})$
Independent Set Lower Bounds

- Wilber I and Wilber II fall in the (new) class of independent rectangle bounds

\[ \text{Theorem: } \text{MinASS}(S) = \Omega(\text{largest independent rectangle set}) \]

- What is the best lower bound in this class?
SignedGreedy

Max of:

- Just fix empty positive-slope rectangles
- Just fix empty negative-slope rectangles

original points
added points

original points
added points

original points
added points
Lower Bounds: SignedGreedy

- **Theorem**: SignedGreedy(S) is within a constant factor of the best independent rectangle bound.

- **Corollary**:
  - $\text{OPT}(S) \geq \text{SignedGreedy}(S)$
  - $\text{SignedGreedy}(S) = \Omega(\text{Wilber I} + \text{Wilber II})$

- SignedGreedy (**lower bound**) is annoyingly similar to GreedyASS (**upper bound**).
Open Problems

- Does GreedyASS share all the nice properties of splay trees?
  - **Recent result**: [Iacono & Pătrașcu] GreedyASS satisfies access theorem (from splay trees)
    - working-set bound
    - entropy bound
    - static finger bound
    - $O(lg n)$ amortized per operation
  - Anyone for dynamic finger?
Open Problems

- Hardness of approximability
- NP-hardness of general position
P.S.

- **ASS** = Arborally Satisfied Set

- **[Arboreal = arboreal = related to trees]**