Optimal Bounds for Open Addressing Without Reordering

Martin Farach-Colton , Andrew Krapivin , William Kuszmaul

Presented by - Amit Roy

March 27, 2025

Theorem - Non-greedy Open-Addressing

Let $n \in \mathbb{N}$ and $\delta \in (0, 1)$ be parameters such that $\delta > \mathcal{O}(1/n)$. There exists an open-addressing hash table that supports $n - \lfloor \delta n \rfloor$ insertions in an array of size n, that does not reorder items after they are inserted, and that offers

- amortized expected probe complexity $O(1)$
- worst-case expected probe complexity $\mathcal{O}(log \delta^{-1})$, and
- worst-case expected insertion time $\mathcal{O}(log \delta^{-1})$.

Theorem - Greedy Open-Addressing

Let $n \in \mathbb{N}$ and $\delta \in (0,1)$ be parameters such that $\delta > \mathcal{O}(1/n^{o(1)})$. There exists a greedy open-addressing strategy that supports $n - \lfloor \delta n \rfloor$ insertions that has

- worst-case expected probe complexity (and insertion time) - $\mathcal{O}(\log^2 \delta^{-1})$
- worst-case probe complexity over all insertions - $\mathcal{O}(\log^2 \delta^{-1} + \log \log n)$, with prob $1 - 1/poly(n)$,
- amortized expected probe complexity - $\mathcal{O}(\log \delta^{-1})$

Lemma

For a given $i \in \alpha$, we have with probability $1 - \frac{1}{n^{\omega(1)}}$ that, after $2|A_i|$ insertion attempts have been made in A_i, fewer than $\frac{\delta}{64}|A_i|$ slots in A_i remain unfilled.

Lemma

The number of keys inserted into $A_{\alpha+1}$ is fewer than $\frac{\delta}{8}n$, with probability $1 - \frac{1}{n^{\omega(1)}}$.

Lemma: Power of two choices

If m balls are placed into n bins by choosing two bins uniformly at random for each ball and placing the ball into the emptier of the two bins, then the maximum load of any bin is $m/n + \log\log n + \mathcal{O}(1)$ with high probability in n.

Algorithm for special array $A_{\alpha+1}$

Algorithm

Input: $n \geq 0$
Output: $y = x^n$
$y \leftarrow 1$
$X \leftarrow x$
$N \leftarrow n$
while $N \neq 0$ **do**
  if N is even **then**
    $X \leftarrow X \times X$
    $N \leftarrow \frac{N}{2}$
  else
    $y \leftarrow y \times X$
    $N \leftarrow N - 1$
  end if
end while

List of Publications

1. Amit Roy and Jayalal Sarma. On Alternation, VC-dimension and k-fold Union of Sets. Proceedings in 11th *European Conference on Combinatorics, Graph Theory and Applications* (EUROCOMB 2021), Sep 6-10, 2021, Barcelona

Thoughts and Questions ?

Thanks

A. Blumer, A. Ehrenfeucht, D. Haussler, and Manfred K. Warmuth.
Learnability and the vapnik-chervonenkis dimension.
J. ACM, 36(4):929–965, October 1989.

Ariel D Procaccia and Jeffrey Rosenschein.
Exact vc-dimension of monotone formulas.
Neural Information Processing -Letters and Reviews, 10, 08 2006.

Eduard Eiben, Fedor V. Fomin, Petr A. Golovach, Willian Lochet, Fahad Panolan, and Kirill Simonov.
Eptas for k-means clustering of affine subspaces.
In *Proceedings of the Thirty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '21, page 2649–2659, USA, 2021. Society for Industrial and Applied Mathematics.

Andrzej Ehrenfeucht, David Haussler, Michael Kearns, and Leslie Valiant.
A general lower bound on the number of examples needed for learning.

Information and Computation, 82(3):247 – 261, 1989.

Tamás Mészáros and Lajos Rónyai.
Shattering-extremal set systems of small vc-dimension.
International Scholarly Research Notices, 2013, 2013.

Thomas Natschläger and Michael Schmitt.
Exact vc-dimension of boolean monomials.
Information Processing Letters, 59(1):19 – 20, 1996.

Ronald L. Rivest.
Learning decision lists.
Machine Learning, 2(3):229–246, Nov 1987.

N Sauer.
On the density of families of sets.
Journal of Combinatorial Theory, Series A, 13(1):145–147, 1972.