Chapter 10

Randomized Algorithms II

By Sariel Har-Peled, December 7, 2009[①] Version: 0.1

10.1 QuickSort with High Probability

One can think about **QuickSort** as playing a game in rounds. Every round, **QuickSort** picks a pivot, splits the problem into two subproblems, and continue playing the game recursively on both subproblems.

If we track a single element in the input, we see a sequence of rounds that involve this element. The game ends, when this element find itself alone in the round (i.e., the subproblem is to sort a single element).

Thus, to show that **QuickSort** takes $O(n \log n)$ time, it is enough to show, that every element in the input, participates in at most $32 \ln n$ rounds with high enough probability.

Indeed, let X_i be the event that the ith element participates in more than $32 \ln n$ rounds.

Let C_{QS} be the number of comparisons performed by **QuickSort**. A comparison between a pivot and an element will be always charged to the element. And as such, the number of comparisons overall performed by **QuickSort** is bounded by $\sum_i r_i$, where r_i is the number of rounds the ith element participated in (the last round where it was a pivot is ignored). We have that

$$\alpha = \mathbf{Pr}[C_{QS} \geq 32 n \ln n] \leq \mathbf{Pr}\left[\bigcup_i X_i\right] \leq \sum_{i=1}^{n} \mathbf{Pr}[X_i] .$$

Here, we used the ***union rule***, that states that for any two events A and B, we have that $\mathbf{Pr}[A \cup B] \leq \mathbf{Pr}[A] + \mathbf{Pr}[B]$. Assume, for the time being, that $\mathbf{Pr}[X_i] \leq 1/n^3$. This implies that

$$\alpha \leq \sum_{i=1}^{n} \mathbf{Pr}[X_i] \leq \sum_{i=1}^{n} 1/n^3 = \frac{1}{n^2}.$$

Namely, **QuickSort** performs at most $32 n \ln n$ comparisons with high probability. It follows, that **QuickSort** runs in $O(n \log n)$ time, with high probability, since the running time of **QuickSort** is proportional to the number of comparisons it performs.

[①]This work is licensed under the Creative Commons Attribution-Noncommercial 3.0 License. To view a copy of this license, visit `http://creativecommons.org/licenses/by-nc/3.0/` or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

To this end, we need to prove that $\mathbf{Pr}[X_i] \leq 1/n^3$.

10.1.1 Proving that an elements participates in small number of rounds.

Consider a run of **QuickSort** for an input made out of n numbers. Consider a specific element x in this input, and let $S_1, S_2, \ldots$ be the subsets of the input that are in the recursive calls that include the element x. Here S_j is the set of numbers in the jth round (i.e., this is the recursive call at depth j which includes x among the numbers it needs to sort).

The element x would be considered to be ***lucky***, in the jth iteration, if the call to the **QuickSort**, splits the current set S_j into two parts, where both parts contains at most $(3/4)\left|S_j\right|$ of the elements.

Let Y_j be an indicator variable which is 1 iff x is lucky in jth round. Formally, $Y_j = 1$ iff $\left|S_j\right|/4 \leq \left|S_{j+1}\right| \leq 3\left|S_j\right|/4$. By definition, we have that

$$\mathbf{Pr}\Big[Y_j\Big] = \frac{1}{2}.$$

Furthermore, $Y_1, Y_2, \ldots, Y_m$ are all independent variables.

Note, that x can participate in at most

$$\rho = \log_{4/3} n \leq 3.5 \ln n \tag{10.1}$$

rounds, since at each successful round, the number of elements in the subproblem shrinks by at least a factor $3/4$, and $|S_1| = n$. As such, if there are ρ successful rounds in the first k rounds, then $|S_k| \leq (3/4)^\rho n \leq 1$.

Thus, the question of how many rounds x participates in, boils down to how many coin flips one need to perform till one gets ρ heads. Of course, in expectation, we need to do this 2ρ times. But what if we want a bound that holds with high probability, how many rounds are needed then?

In the following, we require the following lemma, which we will prove in Section 10.2.

Lemma 10.1.1 *In a sequence of M coin flips, the probability that the number of ones is smaller than $L \leq M/4$ is at most* $\exp(-M/8)$.

To use Lemma 10.1.1, we set

$$M = 32 \ln n \geq 8\rho,$$

see Eq. (10.1). Let Y_j be the variable which is one if x is lucky in the jth level of recursion, and zero otherwise. We have that $\mathbf{Pr}\Big[Y_j = 0\Big] = \mathbf{Pr}\Big[Y_j = 1\Big] = 1/2$ and that $Y_1, Y_2, \ldots, Y_M$ are independent. By Lemma 10.1.1, we have that the probability that there are only $\rho \leq M/4$ ones in $Y_1, \ldots, Y_M$, is smaller than

$$\exp\left(-\frac{M}{8}\right) \leq \exp(-\rho) \leq \frac{1}{n^3}.$$

We have that the probability that x participates in M recursive calls of **QuickSort** to be at most $1/n^3$.

There are n input elements. Thus, the probability that depth of the recursion in **QuickSort** exceeds $32 \ln n$ is smaller than $(1/n^3) * n = 1/n^2$. We thus established the following result.

Theorem 10.1.2 *With high probability (i.e., $1 - 1/n^2$) the depth of the recursion of* **QuickSort** *is $\leq 32 \ln n$. Thus, with high probability, the running time of* **QuickSort** *is $O(n \log n)$.*

Of course, the same result holds for the algorithm **MatchNutsAndBolts** for matching nuts and bolts.

10.2 Chernoff inequality

10.2.1 Preliminaries

Theorem 10.2.1 (*Markov's Inequality*.) *For a non-negative variable X, and $t > 0$, we have:*

$$\mathbf{Pr}[X \geq t] \leq \frac{\mathbf{E}[X]}{t}.$$

Proof: Assume that this is false, and there exists $t_0 > 0$ such that $\mathbf{Pr}[X \geq t_0] > \dfrac{\mathbf{E}[X]}{t_0}$. However,

$$\begin{aligned}
\mathbf{E}[X] &= \sum_x x \cdot \mathbf{Pr}[X = x] \\
&= \sum_{x<t_0} x \cdot \mathbf{Pr}[X = x] + \sum_{x \geq t_0} x \cdot \mathbf{Pr}[X = x] \\
&\geq 0 + t_0 \cdot \mathbf{Pr}[X \geq t_0] \\
&> 0 + t_0 \cdot \frac{\mathbf{E}[X]}{t_0} = \mathbf{E}[X],
\end{aligned}$$

a contradiction. ∎

We remind the reader that two random variables X and Y are *independent* if for any x, y we have that

$$\mathbf{Pr}\big[(X = x) \cap (Y = y)\big] = \mathbf{Pr}[X = x] \cdot \mathbf{Pr}[Y = y].$$

The following claim is easy to verify, and we omit the easy proof.

Claim 10.2.2 *If X and Y are independent, then $\mathbf{E}[XY] = \mathbf{E}[X]\,\mathbf{E}[Y]$.*

If X and Y are independent then $Z = e^X$ and $W = e^Y$ are also independent variables.

10.2.2 Chernoff inequality

Theorem 10.2.3 (*Chernoff inequality*.) *Let $X_1, \ldots, X_n$ be n independent random variables, such that $\mathbf{Pr}[X_i = 1] = \mathbf{Pr}[X_i = -1] = \frac{1}{2}$, for $i = 1, \ldots, n$. Let $Y = \sum_{i=1}^{n} X_i$. Then, for any $\Delta > 0$, we have*

$$\mathbf{Pr}[Y \geq \Delta] \leq \exp\!\left(-\Delta^2/2n\right).$$

Proof: Clearly, for an arbitrary t, to be specified shortly, we have

$$\mathbf{Pr}[Y \geq \Delta] = \mathbf{Pr}\Big[tY \geq t\Delta\Big] = \mathbf{Pr}\Big[\exp(tY) \geq \exp(t\Delta)\Big] \leq \frac{\mathbf{E}\Big[\exp(tY)\Big]}{\exp(t\Delta)}, \tag{10.2}$$

where the first part follows since $\exp(\cdot)$ preserve ordering, and the second part follows by Markov's inequality (Theorem 10.2.1).

Observe that, by the definition of $\mathbf{E}[\cdot]$ and by the Taylor expansion of $\exp(\cdot)$, we have

$$\begin{aligned}
\mathbf{E}\Big[\exp(tX_i)\Big] &= \frac{1}{2}e^t + \frac{1}{2}e^{-t} = \frac{e^t + e^{-t}}{2} \\
&= \frac{1}{2}\left(1 + \frac{t}{1!} + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots\right) \\
&\quad + \frac{1}{2}\left(1 - \frac{t}{1!} + \frac{t^2}{2!} - \frac{t^3}{3!} + \cdots\right) \\
&= \left(1 + \frac{t^2}{2!} + + \cdots + \frac{t^{2k}}{(2k)!} + \cdots\right).
\end{aligned}$$

Now, $(2k)! = k!(k+1)(k+2)\cdots 2k \geq k!2^k$, and thus

$$\mathbf{E}\Big[\exp(tX_i)\Big] = \sum_{i=0}^{\infty} \frac{t^{2i}}{(2i)!} \leq \sum_{i=0}^{\infty} \frac{t^{2i}}{2^i(i!)} = \sum_{i=0}^{\infty} \frac{1}{i!}\left(\frac{t^2}{2}\right)^i = \exp\left(\frac{t^2}{2}\right),$$

again, by the Taylor expansion of $\exp(\cdot)$. Next, by the independence of the X_is, we have

$$\begin{aligned}
\mathbf{E}\Big[\exp(tY)\Big] &= \mathbf{E}\left[\exp\left(\sum_i tX_i\right)\right] = \mathbf{E}\left[\prod_i \exp(tX_i)\right] = \prod_{i=1}^{n} \mathbf{E}\Big[\exp(tX_i)\Big] \\
&\leq \prod_{i=1}^{n} \exp\left(\frac{t^2}{2}\right) = \exp\left(\frac{nt^2}{2}\right).
\end{aligned}$$

We have, by Eq. (10.2), that

$$\mathbf{Pr}[Y \geq \Delta] \leq \frac{\mathbf{E}\Big[\exp(tY)\Big]}{\exp(t\Delta)} \leq \frac{\exp\left(\frac{nt^2}{2}\right)}{\exp(t\Delta)} = \exp\left(\frac{nt^2}{2} - t\Delta\right).$$

Next, we select the value of t that minimizes the right term in the above inequality. Easy calculation shows that the right value is $t = \Delta/n$. We conclude that

$$\mathbf{Pr}[Y \geq \Delta] \leq \exp\left(\frac{n}{2}\left(\frac{\Delta}{n}\right)^2 - \frac{\Delta}{n}\Delta\right) = \exp\left(-\frac{\Delta^2}{2n}\right).$$

■

Note, the above theorem states is that

$$\mathbf{Pr}[Y \geq \Delta] = \sum_{i=\Delta}^{n} \mathbf{Pr}[Y = i] = \sum_{i=n/2+\Delta/2}^{n} \frac{\binom{n}{i}}{2^n} \leq \exp\left(-\frac{\Delta^2}{2n}\right),$$

since $Y = \Delta$ means that we got $n/2 + \Delta/2$ times $+1$s and $n/2 - \Delta/2$ times (-1)s.

By the symmetry of Y, we get the following corollary.

Corollary 10.2.4 *Let $X_1, \ldots, X_n$ be n independent random variables, such that $\mathbf{Pr}[X_i = 1] = \mathbf{Pr}[X_i = -1] = \frac{1}{2}$, for $i = 1, \ldots, n$. Let $Y = \sum_{i=1}^{n} X_i$. Then, for any $\Delta > 0$, we have*

$$\mathbf{Pr}[|Y| \geq \Delta] \leq 2\exp\left(-\frac{\Delta^2}{2n}\right).$$

By easy manipulation, we get the following result.

Corollary 10.2.5 *Let $X_1, \ldots, X_n$ be n independent coin flips, such that $\mathbf{Pr}[X_i = 1] = \mathbf{Pr}[X_i = 0] = \frac{1}{2}$, for $i = 1, \ldots, n$. Let $Y = \sum_{i=1}^{n} X_i$. Then, for any $\Delta > 0$, we have*

$$\mathbf{Pr}\left[\frac{n}{2} - Y \geq \Delta\right] \leq \exp\left(-\frac{2\Delta^2}{n}\right) \quad \textit{and} \quad \mathbf{Pr}\left[Y - \frac{n}{2} \geq \Delta\right] \leq \exp\left(-\frac{2\Delta^2}{n}\right).$$

In particular, we have $\mathbf{Pr}\left[\left|Y - \frac{n}{2}\right| \geq \Delta\right] \leq 2\exp\left(-\frac{2\Delta^2}{n}\right)$.

Proof: Transform X_i into the random variable $Z_i = 2X_i - 1$, and now use Theorem 10.2.3 on the new random variables $Z_1, \ldots, Z_n$. ■

Lemma 10.1.1 (Restatement.) *In a sequence of M coin flips, the probability that the number of ones is smaller than $L \leq M/4$ is at most* $\exp(-M/8)$.

Proof: Let $Y = \sum_{i=1}^{m} X_i$ the sum of the M coin flips. By the above corollary, we have:

$$\mathbf{Pr}[Y \leq L] = \mathbf{Pr}\left[\frac{M}{2} - Y \geq \frac{M}{2} - L\right] = \mathbf{Pr}\left[\frac{M}{2} - Y \geq \Delta\right],$$

where $\Delta = M/2 - L \geq M/4$. Using the above Chernoff inequality, we get

$$\mathbf{Pr}[Y \leq L] \leq \exp\left(-\frac{2\Delta^2}{M}\right) \leq \exp(-M/8).$$

■

10.2.2.1 The Chernoff Bound — General Case

Here we present the Chernoff bound in a more general settings.

Problem 10.2.6 Let $X_1, \ldots X_n$ be n independent Bernoulli trials, where

$$\mathbf{Pr}[X_i = 1] = p_i \ \text{ and } \ \mathbf{Pr}[X_i = 0] = 1 - p_i$$

$$Y = \sum_i X_i \quad \mu = \mathbf{E}[Y].$$

Question: what is the probability that $Y \geq (1 + \delta)\mu$.

Theorem 10.2.7 (Chernoff inequality) *For any $\delta > 0$,*

$$\mathbf{Pr}[Y > (1 + \delta)\mu] < \left(\frac{e^\delta}{(1 + \delta)^{1+\delta}}\right)^\mu.$$

Or in a more simplified form, for any $\delta \leq 2e - 1$,

$$\mathbf{Pr}[Y > (1 + \delta)\mu] < \exp\left(-\mu\delta^2/4\right), \tag{10.3}$$

and

$$\mathbf{Pr}[Y > (1 + \delta)\mu] < 2^{-\mu(1+\delta)},$$

for $\delta \geq 2e - 1$.

Theorem 10.2.8 *Under the same assumptions as the theorem above, we have*

$$\mathbf{Pr}[Y < (1-\delta)\mu] \le \exp\left(-\mu\frac{\delta^2}{2}\right).$$

The proofs of those more general form, follows the proofs shown above, and are omitted. The interested reader can get the proofs from:

http://www.uiuc.edu/~sariel/teach/2002/a/notes/07_chernoff.ps