ISSN:2249-5789
Amitesh Singh et al , International Journal of Computer Science & Communication Networks,Vol 2(4), 473-474

Asymptotic Analysis of Randomized Quick Select using Generating function

Amitesh Singh
singh.amitesh@gmail.com
Anuj Agarwal
anuj.ism@gmail.com

Abstract

QUICK-SELECT (modelled after the quicksort algorithm) can be used to find the i^{th} order statistics (smallest or largest) out of a set of n elements. It is also known as Hoare's FIND algorithm[7]. In this paper, we represent standard version, where the pivot is chosen from the given input of n elements. Knuth[1] has already calculated the average cost, by elementary method.In this paper, we find the explicit formulae for the average cost by using bivariate generating function and differential equations solving techniques.

1 Introduction

The problem of *select* consists of finding the i^{th} smallest (or i^{th} largest) element out of a given set of n elements.
The following pseudo-code for QUICK-SELECT returns the i^{th} smallest element of the array A[p..r],$1 \leq i \leq n$.

```
QUICK-SELECT(A,p,r,i)
if p=r then
|  then return A[p]
end
q ← RANDOMIZED-PARTITION(A,p,r)
k ← q-p+1
if i = k then
|  return A[q]
else if i < k then
|  return QUICK-SELECT(A,p,q-1,i)
else
|  return QUICK-SELECT(A,q+1,r,i-k)
```

Algorithm 1: Randomized Select

2 The Average Cost

Procedure RANDOMIZED-PARTITION is likely to return an element as pivot with probability that we denote by $P_{n,k}$, so we have

$$P_{n,k} = \frac{1}{n} \tag{1}$$

The average cost by QUICK-SELECT on an input array $A[p..r]$ of n elements is a random variable that we denote by $R_{n,i}$.

$$R_{n,i} = n - 1 + C + \sum_{k=1}^{i-1} P_{n,k}\ R_{n-k,i-k} + \sum_{k=i+1}^{n} P_{n,k}\ R_{k-1,i}$$

and $R_{n,i} = R_{n,n+1-i}$ for any i, $1 \leq i \leq n$
where C is the cost involved in finding a pivot.

For convenience, we took $C \approx 0$.

$$R_{n,i} = n - 1 + \sum_{k=1}^{i-1} P_{n,k}\ R_{n-k,i-k} + \sum_{k=i+1}^{n} P_{n,k}\ R_{k-1,i} \tag{2}$$

Lets set up the generating function.

$$R_i(x) = \sum_{n \geq i} R_{n,i}\ x^n \tag{3}$$

Multiply the equation (2) by $\binom{n}{1}\ x^{n-1}$ and summing up, we obtain

$$R_i^{'}(x)$$

$$= \sum_{n \geq i} n(n-1)\ x^{n-1} \tag{4}$$

$$+ \sum_{n \geq i} \sum_{k=1}^{i-1} R_{n-k,i-k}\ x^{n-1} \tag{5}$$

$$+ \sum_{n \geq i} \sum_{k=i+1}^{n} R_{k-1,i}\ x^{n-1} \tag{6}$$

Solving each part separately,
Equation (4)

$$\sum_{n \geq i} n(n-1)\ x^{n-1} = \sum_{n \geq 1} n(n-1)\ x^{n-1} - \sum_{n \geq 1}^{i-1} n(n-1)\ x^{n-1}$$

$$= \frac{2x}{(1-x)^3} - \sum_{k \geq 1}^{i-1} k(k-1)\ x^{k-1}$$

Equation (5)

Since,

$$R_{i-k}(x) = \sum_{n-k \geq i-k} R_{n-k,i-k} x^{n-k}$$

[by 3]

$$R_{i-k}(x) = \sum_{n \geq i} R_{n-k,i-k} x^{n-k}$$

Hence,

$$\sum_{n \geq i} \sum_{k=1}^{i-1} R_{n-k,i-k}\ x^{n-1} = \sum_{k=1}^{i-1} x^{k-1} R_{i-k}(x)$$

Equation (6)

$$\sum_{n\geq i}\sum_{k=i+1}^{n} R_{k-1,i}\, x^{n-1} = \sum_{k\geq i+1} R_{k-1,i}\, x^{k-1} \sum_{n\geq k} x^{n-k}$$
$$= \sum_{k\geq i+1} R_{k-1,i}\, x^{k-1} \frac{1}{(1-x)} = \frac{R_i(x)}{(1-x)}$$

Hence,

$$R_i^{'}(x) = \frac{2x}{(1-x)^3} - \sum_{k\geq 1}^{i-1} k(k-1)\, x^{k-1} + \sum_{k=1}^{i-1} x^{k-1} R_{i-k}(x) + \frac{R_i(x)}{(1-x)} \quad (7)$$

This equation is valid for $i \geq 1$ and the initial values are $R_i(0) = 0$.
Lets define the double generating function.

$$R(x,y) := \sum_{i\geq 1} R_i(x) y^i$$

Multiply the equation (7) by y^j and summing up, we obtain

$$\sum_{i\geq 1} R_i^{'}(x)\, y^i = \frac{2x}{(1-x)^3} \sum_{i\geq 1} y^i - \sum_{i\geq 1}\sum_{k\geq 1}^{i-1} k(k-1)\, x^{k-1}\, y^i + \sum_{i\geq 1}\sum_{k=1}^{i-1} x^{k-1} R_{i-k}(x)\, y^i + \frac{1}{(1-x)} \sum_{i\geq 1} R_i(x)\, y^i$$

$$\frac{\partial}{\partial x} R(x,y) = \frac{2xy}{(1-x)^3(1-y)} - xy^3 \sum_{k\geq 1} k(k-1)(xy)^{k-2} \sum_{i-k\geq 1} y^{i-k-1} + y\sum_{k\geq 1}(xy)^{k-1} \sum_{i-k\geq 1} R_{i-k}(x)\, y^{i-k} + \frac{R(x,y)}{(1-x)}$$

$$\frac{\partial}{\partial x} R(x,y) = \frac{2xy}{(1-x)^3(1-y)} - \frac{2xy^3}{(1-xy)^3(1-y)} \frac{yR(x,y)}{(1-xy)} + \frac{R(x,y)}{(1-x)}$$

$$\frac{\partial}{\partial x} R(x,y) - R(x,y)\left\{\frac{y}{(1-xy)} + \frac{1}{(1-x)}\right\} = \frac{2xy}{1-y}\left\{\frac{1}{(1-x)^3} - \frac{y^2}{(1-xy)^3}\right\} \quad (8)$$

with initial conditions $R(0,y) = 0$.
Solving equation (8)

$$R(x,y) = 2\Big(\frac{xy}{(1-x)(1-y)} + \frac{y}{(1-x)^2(1-xy)} + \frac{y(2y-1)\log(\frac{1}{1-x})}{(1-y)(1-x)(1-xy)} + \frac{(y-2)\log(\frac{1}{1-xy})}{(1-x)(1-y)(1-xy)} + \frac{1}{(1-x)(1-yx)^2} - \frac{(y+1)}{(1-x)(1-xy)}\Big) \quad (9)$$

We have to find the coefficient of $[x^n\ y^i]$ of $R(x,y)$ i.e. $R_{n,i} = [x^n y^i]R(x,y)$

$[x^n y^i]\frac{yx}{(1-x)(1-xy)} = 1$, $[x^n y^i]\frac{y}{(1-x)^2(1-xy)} = n-i+2$,
$[x^n y^i]\frac{\log(\frac{1}{1-x})}{(1-x)(1-y)(1-xy)} = (n+1)(H_{n+1}-1)-(n-i)(H_{n-i}-1)$,
$[x^n y^i]\frac{y(2y-1)\log(\frac{1}{1-x})}{(1-x)(1-y)(1-xy)} = (n+1)(H_{n+1}-1)-(n-i+3)H_{n-i+1}-i$
$[x^n y^i]\frac{\log(\frac{1}{1-xy})}{(1-x)(1-y)(1-xy)} = (i+1)(H_{i+1}-1)$,
$[x^n y^i]\frac{(y-2)\log(\frac{1}{1-xy})}{(1-x)(1-y)(1-xy)} = -(i+2)H_i + i$
$[x^n y^i]\frac{1}{(1-x)(1-xy)^2} = i+1$, $[x^n y^i]\frac{1}{(1-x)(1-xy)} = 1$

We find the well known result. [1]

$$R(x,y) = 2(n+3+(n+1)H_n - (i+2)H_i - (n-i+3)H_{n-i+1}) \quad (10)$$

where $H_n = \sum_{k=1}^{n}\frac{1}{k}$ is the n^{th} harmonic number.

3 Summary

We get the running time of Quick Select in average case is $O(n)$ as expected (10). The method of using generating function can be extended to solve quick select recurrence where selection of pivot is done by choosing median of $2t+1$ elements.

References

[1] Donald E Knuth, *The Art of Computer Programming - Volume 3, Second Edition*, PEARSON, 2011.

[2] Ronald L. Graham, Donald E. Knuth, Oren Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, PEARSON, 2011.

[3] Donald E Knuth, *The Art of Computer Programming - Volume 1, Second Edition*, PEARSON, 2011.

[4] Robert Sedgewick, Philippe Flajolet, *An Introduction to the Analysis of Algorithms*, PEARSON, 2010.

[5] Conrado Martinez, *Partial Quicksort.*

[6] Thomas H. Cormen, C. E. Leiserson,Ronald L. Rivest, Clifford Stein, *Introduction to Algorithms, Second Edition*, PEARSON, 2010.

[7] C. A. R. Hoare, *Find Algorithm 65*, ACM,1961.