Quick Sort

Prof. Micha Hofri

March 9, 2014

First let's take a look at a version of the Quick Sort Algorithm:

```
Data: A, i, j
QS(A, i, j);
if i > j then
|  return;
end
k=partition(A, i, j);
QS(A, i, k − 1);
QS(A, k + 1, j);
```

Algorithm 1: One Version of Quick Sort Algorithm

In the algorithm, input A is an array with random permutation of size n. All the entries are with distinctive value.
Apparently, the major part of the QS algorithm is the partition part, now let's look at a version of the partition algorithm.

```
partition(A, lo, hi);
p = A[hi];
i = lo-1;
for i=lo to hi do
|  if A[j] <= p then
|  |  i++;
|  |  swap(A[i],A[j]);
|  end
end
swap(A[i+1], A[hi]);
return i+1
```

Algorithm 2: A Version of Partition

Define V_n as the number of calls to QS, hence we have:

$$V_0 = 1, V_1 = 3$$

$$V_n = 1 + V_{k-1} + V_{n-k}$$

Let a subset of $\lfloor b \rfloor$ values in $[n]$ be observed in all $n!$ permutations, then the number of possibilities is:

$$\binom{n}{b}(n-b)!$$

Because first, we choose b terms from n, which is $\binom{n}{b}$. And then the $(n-b)$ numbers are permutated.
Define t as the number of terms on the right side not moved; p as the value of the pivot(is p means this?).
Then we have:

$$\frac{1}{t!}\cdot\frac{1}{(n-p-t)!}\cdot\frac{1}{\binom{n-p}{t}}=\frac{1}{(n-p)!}$$

(I forgot what is the above function means)
Next, let's compute the value of $E[V_n]$:

$$\begin{aligned}
E[V_n] &= v_n \\
&= 1+\sum_{n=1}^{n} P[K=k](v_{k-1}+v_{n-k}) \\
&= 1+\frac{1}{n}(\sum_{k=1}^{n} v_{k-1}+\sum_{k=1}^{n} v_{n-k}) \\
&= 1+\frac{2}{n}\sum_{j=0}^{n-1} v_j \\
nv_n &= n+2\sum_{j=0}^{n-1} v_j \\
(n+1)v_{n+1} &= n+1+2\sum_{j=0}^{n} v_j
\end{aligned}$$

Substract the above two equations, we get:

$$\begin{aligned}
(n+1)v_{n+1}-nv_n &= 1+2v_n \\
(n+1)v_{n+1} &= 1+(n+2)v_n \\
\frac{v_{n+1}}{n+2} &= \frac{1}{(n+1)(n+2)}+\frac{v_n}{n+1}
\end{aligned}$$

Let $u_n=\frac{v_n}{n+1}$

$$
\begin{aligned}
u_{n+1} &= \frac{1}{(n+1)(n+2)} + u_n \\
u_{n+1} &= \frac{3}{2} + \sum_{i=2}^{n+1} \frac{1}{i(i+1)} \\
&= \frac{3}{2} + \frac{1}{2} - \frac{1}{i+2} \\
&= 2 - \frac{1}{n+2} \\
\frac{v_{n+1}}{n+2} &= 2 - \frac{1}{n+2} \\
v_{n+1} &= 2(n+2) - 1 \\
&= 2n+3 \\
v_n &= 2(n-1)+3 \\
&= 2n+1
\end{aligned}
$$

Finally, we get:

$E[V_n] = 2n+1$

Now, let's look at another version of the QS algorithm, which doesn't compare the value of i and j at the beginning.

```
QS(A, i, j);
if i < k − 1 then
 |  QS(A, i, k − 1)
end
if j > k − 1 then
 |  QS(A, k + 1, j)
end
```

Algorithm 3: Another Version of the Quick Sort Algorithm

Agian, we calculate the expected value of V_n: Number of calls to the QS function:

Same as before, assume $v_n = E[V_n]$:

First, we have:

$$v_0 = 0 \qquad v_1 = 0 \qquad v_2 = 1$$

Same as before, we have:

$$
\begin{aligned}
\frac{v_{n+1}}{n+2} &= u_{n+1} \\
&= u_n + \frac{1}{(n+1)(n+2)} \\
&= u_2 + \sum_{k=3}^{n+1} \frac{1}{k(k+1)} \\
&= u_2 = \frac{1}{3} - \frac{1}{n+2}
\end{aligned}
$$

Because $u_2 = \frac{v_2}{3} = \frac{1}{3}$, so

$$
\begin{aligned}
u_{n+1} &= \frac{2}{3} - \frac{1}{n+2} \\
&= \frac{1}{3} \cdot \frac{2n+1}{n+2}
\end{aligned}
$$

hence:

$$
v_n = \frac{2n+1}{3}
$$

Next, let's consider the number of term comparisons.
Define C_n as the number of term comparisons. The term comparison operations all occurs in the partition function. Define

$$
X_{ij} \begin{cases} 1 & \text{if } y_i \text{ and } y_j \text{ are compared} \\ 0 & \text{if o.w.} \end{cases}
$$

Then we have:

$$
P[X_{ij} = 1] = \frac{2}{j - i + 1}
$$

Since at least on of y_i or y_j to be chosen as the pivot to make them compared. Then we have:

$$
C_n = \sum_{1 \le i < j \le n} x_{ij}
$$

So:

$$E[C_n] = \sum_{i+1}^{n} \sum_{j=i+1}^{n} \frac{2}{j-i+1}$$

Let $k = j - i + 1$

$$\begin{aligned}
&= \sum_{i=1}^{n} \sum_{k=2}^{n-i+1} \frac{2}{k} \\
&= \sum_{k=2}^{n-1} \sum_{i=1}^{n-k+1} \frac{2}{k} \\
&= \sum_{k=2}^{n-1} \frac{1}{k} \sum_{i=1}^{n-k+1} 1 \\
&= \sum_{k=2}^{n} \frac{n+1-k}{k} \\
&= (n+1) \sum_{k=2}^{n} \frac{1}{k} - \sum_{k+2}^{n} 1 \\
&= (n+1)(H_n - 1) - n + 1 \\
&= 2((n+1)H_n - 2n) \\
&= 2(n+1)H_n - 4n
\end{aligned}$$

Another Method:

$$\sum_{r=1}^{n} H_r = \sum_{r=1}^{n} \sum_{k=1}^{r} \frac{1}{k} = (n+1)H_n - n$$

Hence:

$$\begin{aligned}
2 \sum_{i=1}^{n} \sum_{k=2}^{n-i+1} &= 2 \sum_{i=1}^{n} (H_{n-i+1} - 1) \\
&= -2n + \sum_{i=1}^{n} H_{n-i+1} \\
&= -2n + 2 \sum_{r=1}^{n} H_r \\
&= 2(n+1)H_n - 4n
\end{aligned}$$