Quick Sort

EUnS

May 5, 2020

```
QUICKSORT(A, p, r)
    if p < r
        q = PARTITION(A, p, r)
        QUICKSORT(A, p, q - 1)
        QUICKSORT(A, q + 1, r)

```

```
QUICKSORT(A, p, r)
    if p < r
        q = PARTITION(A, p, r)
        QUICKSORT(A, p, q - 1)
        QUICKSORT(A, q + 1, r)
```

```
PARTITION(A, p, r) // Lumoto
    x =A[r]
    i = p - 1
    for j = p to r - 1
        if A[j] <= x
            i = i + 1
            exchange A[i] with A[j]
    exchange A[i + 1] with A[r]
    return i + 1
```

최악의 경우 분할 케이스

최악의 경우 분할 케이스

- 왼쪽 오른쪽 분할이 한쪽으로 쏠려(오름차순,내림차순) 극도로 불균형하게 일어났을때

최악의 경우 분할 케이스

- 왼쪽 오른쪽 분할이 한쪽으로 쏠려(오름차순,내림차순) 극도로 불균형하게 일어났을때

$$T(n) = T(n-1) + cn$$

최악의 경우 분할 케이스

- 왼쪽 오른쪽 분할이 한쪽으로 쏠려(오름차순,내림차순) 극도로 불균형하게 일어났을때

$$T(n) = T(n-1) + cn$$

$$T(n) = T(n-1) + cn$$

최악의 경우 분할 케이스

- 왼쪽 오른쪽 분할이 한쪽으로 쏠려(오름차순,내림차순) 극도로 불균형하게 일어났을때

$$T(n) = T(n-1) + cn$$

$$\begin{aligned} T(n) &= T(n-1) + cn \\ &= T(n-2) + c(n-1) + cn \end{aligned}$$

최악의 경우 분할 케이스

- 왼쪽 오른쪽 분할이 한쪽으로 쏠려(오름차순,내림차순) 극도로 불균형하게 일어났을때

$$T(n) = T(n-1) + cn$$

$$\begin{aligned} T(n) &= T(n-1) + cn \\ &= T(n-2) + c(n-1) + cn \\ &= c\sum_{k=1}^{n} k \end{aligned}$$

최악의 경우 분할 케이스

- 왼쪽 오른쪽 분할이 한쪽으로 쏠려(오름차순,내림차순) 극도로 불균형하게 일어났을때

$$T(n) = T(n-1) + cn$$

$$\begin{aligned} T(n) &= T(n-1) + cn \\ &= T(n-2) + c(n-1) + cn \\ &= c\sum_{k=1}^{n} k \\ &= \frac{1}{2}cn^2 \end{aligned}$$

최악의 경우 분할 케이스

- 왼쪽 오른쪽 분할이 한쪽으로 쏠려(오름차순,내림차순) 극도로 불균형하게 일어났을때

$$T(n) = T(n-1) + cn$$

$$\begin{aligned} T(n) &= T(n-1) + cn \\ &= T(n-2) + c(n-1) + cn \\ &= c\sum_{k=1}^{n} k \\ &= \frac{1}{2}cn^2 \\ &= \Theta(n^2) \end{aligned}$$

최선의 분할 케이스

최선의 분할 케이스

- 정확하게 반으로 나누어 졌을때

최선의 분할 케이스

- 정확하게 반으로 나누어 졌을때

$$T(n) = 2T\left(\frac{n}{2}\right) + cn$$

- 마스터 정리 case 2

$$\Theta(n \log n)$$

- 항상 9:1로 분할하는 경우
- 최악의 경우와 최선의 경우가 번갈아 나타나는 경우

9 : 1로 분할

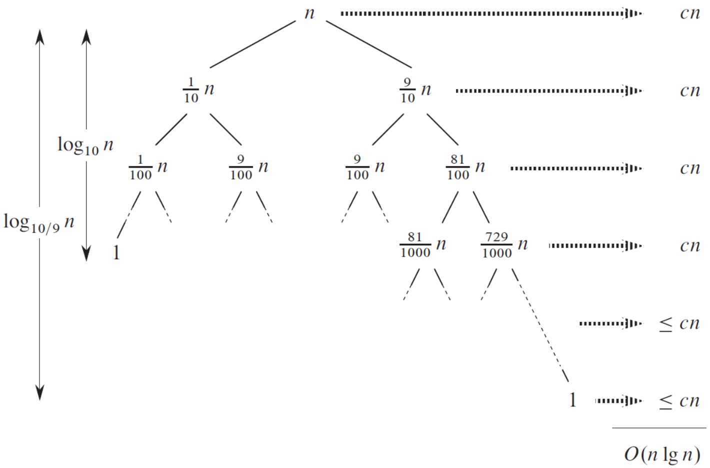

Figure: 9:1로 분할하는 재귀 트리[?]

- $T(n) \leq T\left(\frac{9n}{10}\right) + T\left(\frac{n}{10}\right) + cn$

- $T(n) \leq T\left(\frac{9n}{10}\right) + T\left(\frac{n}{10}\right) + cn$
- $n \log_{10/9} n = O(n \log n)$

최악 최선

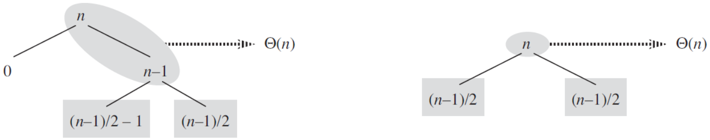

Figure: 최악,최선이 번갈아 나타나는 재귀 트리[?]

최악 최선

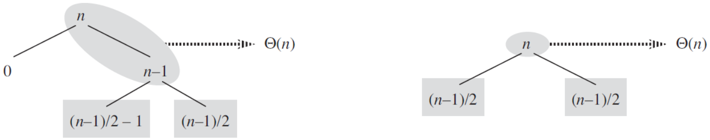

Figure: 최악,최선이 번갈아 나타나는 재귀 트리[**?**]

- $T(n)$ $T(n-1)$ 둘다 $\Theta(n)$

최악 최선

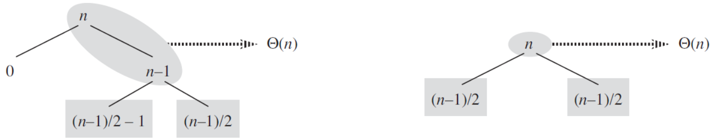

Figure: 최악,최선이 번갈아 나타나는 재귀 트리[?]

- $T(n)$ $T(n-1)$ 둘다 $\Theta(n)$
- $\Theta(n \lg n)$

최악의 케이스 분석

최악의 케이스 분석

- $$T(n) = \max_{0 \leq q \leq n-1} (T(q) + T(n-q-1)) + \Theta(n)$$

최악의 케이스 분석

- $$T(n) = \max_{0 \le q \le n-1} (T(q) + T(n-q-1)) + \Theta(n)$$
- $T(n) \le c_1 n^2$이라 가정

최악의 케이스 분석

- $$T(n) = \max_{0 \le q \le n-1} (T(q) + T(n-q-1)) + \Theta(n)$$
- $T(n) \le c_1 n^2$이라 가정
- $$T(n) \le \max_{0 \le q \le n-1} (c_1 q^2 + c_1 (n-q-1)^2) + \Theta(n)$$

최악의 케이스 분석

- $$T(n) = \max_{0 \le q \le n-1} (T(q) + T(n-q-1)) + \Theta(n)$$
- $T(n) \le c_1 n^2$이라 가정
- $$T(n) \le \max_{0 \le q \le n-1} (c_1 q^2 + c_1 (n-q-1)^2) + \Theta(n)$$
- $f(q) = c_1 q^2 + c_1 (n-q-1)^2 (0 \le q \le n-1)$

최악의 케이스 분석

- $$T(n) = \max_{0 \leq q \leq n-1} (T(q) + T(n-q-1)) + \Theta(n)$$
- $T(n) \leq c_1 n^2$이라 가정
- $$T(n) \leq \max_{0 \leq q \leq n-1} (c_1 q^2 + c_1 (n-q-1)^2) + \Theta(n)$$
- $f(q) = c_1 q^2 + c_1 (n-q-1)^2 (0 \leq q \leq n-1)$
- 극댓값이자 최댓값은 $q = 0$일때와 $q = n-1$일때 둘중하나.

$$T(n) \leq c_1(n-1)^2 + \Theta(n)$$

$$
\begin{aligned}
T(n) &\le c_1(n-1)^2 + \Theta(n) \\
&\le c_1 n^2 - c(2n-1) + \Theta(n)
\end{aligned}
$$

$$
\begin{aligned}
T(n) &\le c_1(n-1)^2 + \Theta(n) \\
&\le c_1 n^2 - c(2n-1) + \Theta(n) \\
&\le c_1 n^2
\end{aligned}
$$

기대 수행 시간

Lemma

크기가 n인 배열에 대해 QUICKSORT의 전체 실행 동안 PARTITION의 4행에서 수행되는 비교 횟수를 X라 하자. QUICKSORT의 실행시간은 O(N+X)이다.

기대 수행 시간

Lemma

크기가 n인 배열에 대해 QUICKSORT의 전체 실행 동안 PARTITION의 4행에서 수행되는 비교 횟수를 X라 하자. QUICKSORT의 실행시간은 O(N+X)이다.

Proof.

알고리즘은 PARTITION을 최대 n회 호출한다. 각 호출은 for 루프를 실행하는데, for 루프의 각 반복은 4행 비교문을 실행한다. □

- 각 호출에 대한 비교를 분석하지않고 총 비교수를 계산

- 각 호출에 대한 비교를 분석하지않고 총 비교수를 계산
- 배열 $A = \{z_1, z_2, ..., z_n\}$의 각 요소가 내림차순으로 정렬

- 각 호출에 대한 비교를 분석하지않고 총 비교수를 계산
- 배열 $A = \{z_1, z_2, ..., z_n\}$의 각 요소가 내림차순으로 정렬
- 집합 $Z_{ij} = \{z_i, z_{i+1}, ..., z_j\}$

- 각 호출에 대한 비교를 분석하지않고 총 비교수를 계산
- 배열 $A = \{z_1, z_2, ..., z_n\}$의 각 요소가 내림차순으로 정렬
- 집합 $Z_{ij} = \{z_i, z_{i+1}, ..., z_j\}$
- z_i와 z_j는 최대 한번 비교

- 각 호출에 대한 비교를 분석하지않고 총 비교수를 계산
- 배열 $A = \{z_1, z_2, ..., z_n\}$의 각 요소가 내림차순으로 정렬
- 집합 $Z_{ij} = \{z_i, z_{i+1}, ..., z_j\}$
- z_i와 z_j는 최대 한번 비교
- PARTITION에서 비교를 하는 경우는 하나의 원소가 pivot으로 선택 되었을 때인데, 이후에 이 pivot은 절대로 다른 원소와 비교하지 않는다.

- $X_{ij} = I\{z_i$가 z_j와 비교한다 $\}$

- $X_{ij} = I\{z_i\text{가 } z_j\text{와 비교한다 }\}$
- $X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}$

$$E[x] = E\left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right]$$

- $X_{ij} = I\{z_i$가 z_j와 비교한다 $\}$
- $X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}$

$$
\begin{aligned}
E[x] &= E\left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right] \\
&= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E\left[X_{ij}\right]
\end{aligned}
$$

- $X_{ij} = I\{z_i\text{가 } z_j\text{와 비교한다 }\}$
- $X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}$

$$\begin{aligned} E[x] &= E\left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right] \\ &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E\left[X_{ij}\right] \\ &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} Pr\{z_i\text{가}z_j\text{와 비교한다}\} \end{aligned}$$

$$Pr\{z_i\text{가}z_j\text{와 비교한다}\}$$

$$
\begin{aligned}
&Pr\{z_i\text{가}z_j\text{와 비교한다}\} \\
&= Pr\{Z_{ij}\text{에서}z_i\text{또는}z_j\text{가 첫번째로 선택된다.}\}
\end{aligned}
$$

$$
\begin{aligned}
&Pr\{z_i\text{가}z_j\text{와 비교한다}\}\\
&= Pr\{Z_{ij}\text{에서}z_i\text{또는}z_j\text{가 첫번째로 선택된다.}\}\\
&= Pr\{Z_{ij}\text{에서}z_i\text{가 첫번째로 선택된다.}\}
\end{aligned}
$$

$$
\begin{aligned}
& Pr\{z_i\text{가}z_j\text{와 비교한다}\} \\
&= Pr\{Z_{ij}\text{에서}z_i\text{또는}z_j\text{가 첫번째로 선택된다.}\} \\
&= Pr\{Z_{ij}\text{에서}z_i\text{가 첫번째로 선택된다.}\} \\
&+ Pr\{Z_{ij}\text{에서}z_j\text{가 첫번째로 선택된다.}\}
\end{aligned}
$$

$$
\begin{aligned}
& Pr\{z_i\text{가}z_j\text{와 비교한다}\} \\
&= Pr\{Z_{ij}\text{에서}z_i\text{또는}z_j\text{가 첫번째로 선택된다.}\} \\
&= Pr\{Z_{ij}\text{에서}z_i\text{가 첫번째로 선택된다.}\} \\
&+ Pr\{Z_{ij}\text{에서}z_j\text{가 첫번째로 선택된다.}\} \\
&= \frac{1}{j-i+1} + \frac{1}{j-i+1}
\end{aligned}
$$

$$
\begin{aligned}
& Pr\{z_i\text{가 } z_j\text{와 비교한다}\} \\
= {} & Pr\{Z_{ij}\text{에서 } z_i \text{ 또는 } z_j\text{가 첫번째로 선택된다.}\} \\
= {} & Pr\{Z_{ij}\text{에서 } z_i\text{가 첫번째로 선택된다.}\} \\
+ {} & Pr\{Z_{ij}\text{에서 } z_j\text{가 첫번째로 선택된다.}\} \\
= {} & \frac{1}{j-i+1} + \frac{1}{j-i+1} \\
= {} & \frac{2}{j-i+1}
\end{aligned}
$$

상한

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

상한

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

$j-i$을 k로 치환

상한

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

$j - i$을 k로 치환

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

상한

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

$j-i$을 k로 치환

$$\begin{aligned} E[X] &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1} \\ &= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k+1} \end{aligned}$$

상한

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

$j - i$을 k로 치환

$$\begin{aligned} E[X] &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1} \\ &= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k+1} \\ &\leq \sum_{i=1}^{n-1} \sum_{k=1}^{n} \frac{2}{k} \end{aligned}$$

상한

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

$j - i$을 k로 치환

$$\begin{aligned} E[X] &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1} \\ &= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k+1} \\ &\le \sum_{i=1}^{n-1} \sum_{k=1}^{n} \frac{2}{k} \\ &\le \sum_{i=1}^{n-1} c \lg n \end{aligned}$$

상한

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

$j - i$을 k로 치환

$$\begin{aligned} E[X] &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1} \\ &= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k+1} \\ &\leq \sum_{i=1}^{n-1} \sum_{k=1}^{n} \frac{2}{k} \\ &\leq \sum_{i=1}^{n-1} c \lg n \\ &\leq cn \lg n \end{aligned}$$

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

$$
\begin{aligned}
E[X] &= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k+1} \\
&= \sum_{k=1}^{n-1} \frac{2}{k+1} + \sum_{k=1}^{n-2} \frac{2}{k+1} + \cdots + \sum_{k=1}^{1} \frac{2}{k+1}
\end{aligned}
$$

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

$$= \sum_{k=1}^{n-1} \frac{2}{k+1} + \sum_{k=1}^{n-2} \frac{2}{k+1} + \cdots + \sum_{k=1}^{1} \frac{2}{k+1}$$

$$= (n-1)\frac{2}{1+1} + (n-2)\frac{2}{2+1} + \cdots + 1 \times \frac{2}{(n-1)+1}$$

$$
\begin{aligned}
E[X] &= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k+1} \\
&= \sum_{k=1}^{n-1} \frac{2}{k+1} + \sum_{k=1}^{n-2} \frac{2}{k+1} + \cdots + \sum_{k=1}^{1} \frac{2}{k+1} \\
&= (n-1)\frac{2}{1+1} + (n-2)\frac{2}{2+1} + \cdots + 1 \times \frac{2}{(n-1)+1} \\
&= \sum_{k=1}^{n-1} \frac{2}{k+1} \times (n-k)
\end{aligned}
$$

$$
\begin{aligned}
E[X] &= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k+1} \\
&= \sum_{k=1}^{n-1} \frac{2}{k+1} + \sum_{k=1}^{n-2} \frac{2}{k+1} + \cdots + \sum_{k=1}^{1} \frac{2}{k+1} \\
&= (n-1)\frac{2}{1+1} + (n-2)\frac{2}{2+1} + \cdots + 1 \times \frac{2}{(n-1)+1} \\
&= \sum_{k=1}^{n-1} \frac{2}{k+1} \times (n-k) \\
&= \sum_{k=1}^{n-1} \left(\frac{2n}{k+1} - \frac{2k}{k+1} \right)
\end{aligned}
$$

$$
\begin{aligned}
E[X] &= \sum_{i=1}^{n-1}\sum_{k=1}^{n-i}\frac{2}{k+1} \\
&= \sum_{k=1}^{n-1}\frac{2}{k+1} + \sum_{k=1}^{n-2}\frac{2}{k+1} + \cdots + \sum_{k=1}^{1}\frac{2}{k+1} \\
&= (n-1)\frac{2}{1+1} + (n-2)\frac{2}{2+1} + \cdots + 1 \times \frac{2}{(n-1)+1} \\
&= \sum_{k=1}^{n-1}\frac{2}{k+1} \times (n-k) \\
&= \sum_{k=1}^{n-1}\left(\frac{2n}{k+1} - \frac{2k}{k+1}\right) \\
&= 2n\sum_{k=1}^{n-1}\frac{1}{k+1} - 2\sum_{k=1}^{n-1}\frac{k}{k+1}
\end{aligned}
$$

$$
\begin{aligned}
E[X] &= \sum_{i=1}^{n-1}\sum_{k=1}^{n-i}\frac{2}{k+1} \\
&= \sum_{k=1}^{n-1}\frac{2}{k+1} + \sum_{k=1}^{n-2}\frac{2}{k+1} + \cdots + \sum_{k=1}^{1}\frac{2}{k+1} \\
&= (n-1)\frac{2}{1+1} + (n-2)\frac{2}{2+1} + \cdots + 1 \times \frac{2}{(n-1)+1} \\
&= \sum_{k=1}^{n-1}\frac{2}{k+1} \times (n-k) \\
&= \sum_{k=1}^{n-1}\left(\frac{2n}{k+1} - \frac{2k}{k+1}\right) \\
&= 2n\sum_{k=1}^{n-1}\frac{1}{k+1} - 2\sum_{k=1}^{n-1}\frac{k}{k+1} \\
&\geq 2nc\lg n - 2(n-1)\left(\because -\sum_{k=1}^{n-1}\frac{k}{k+1} \geq -\sum_{k=1}^{n-1}\left(\frac{k}{k+1}+\frac{1}{k+1}\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
E[X] &= \sum_{i=1}^{n-1}\sum_{k=1}^{n-i}\frac{2}{k+1} \\
&= \sum_{k=1}^{n-1}\frac{2}{k+1} + \sum_{k=1}^{n-2}\frac{2}{k+1} + \cdots + \sum_{k=1}^{1}\frac{2}{k+1} \\
&= (n-1)\frac{2}{1+1} + (n-2)\frac{2}{2+1} + \cdots + 1 \times \frac{2}{(n-1)+1} \\
&= \sum_{k=1}^{n-1}\frac{2}{k+1} \times (n-k) \\
&= \sum_{k=1}^{n-1}\left(\frac{2n}{k+1} - \frac{2k}{k+1}\right) \\
&= 2n\sum_{k=1}^{n-1}\frac{1}{k+1} - 2\sum_{k=1}^{n-1}\frac{k}{k+1} \\
&\geq 2nc\lg n - 2(n-1)\left(\because -\sum_{k=1}^{n-1}\frac{k}{k+1} \geq -\sum_{k=1}^{n-1}\left(\frac{k}{k+1}+\frac{1}{k+1}\right)\right) \\
&\geq \Omega(n\lg n)
\end{aligned}
$$

따라서

$$\Theta(n \lg n)$$