Lower Bound For Comparison-Based Sorting

Note Title 10/25/2007

1. What is the upper bound for comparison-based sorting?

notes:

quicksort $O(n^2)$ in worst case
insertion sort $O(n)$ in best case

merge sort ~~Θ~~$(n\log n)$ in worst case
upper bound is $n\log n$

2. So what is the lower bound?

Use the same strategy as before -

<u>generic decision tree</u>

Generic Decision Tree

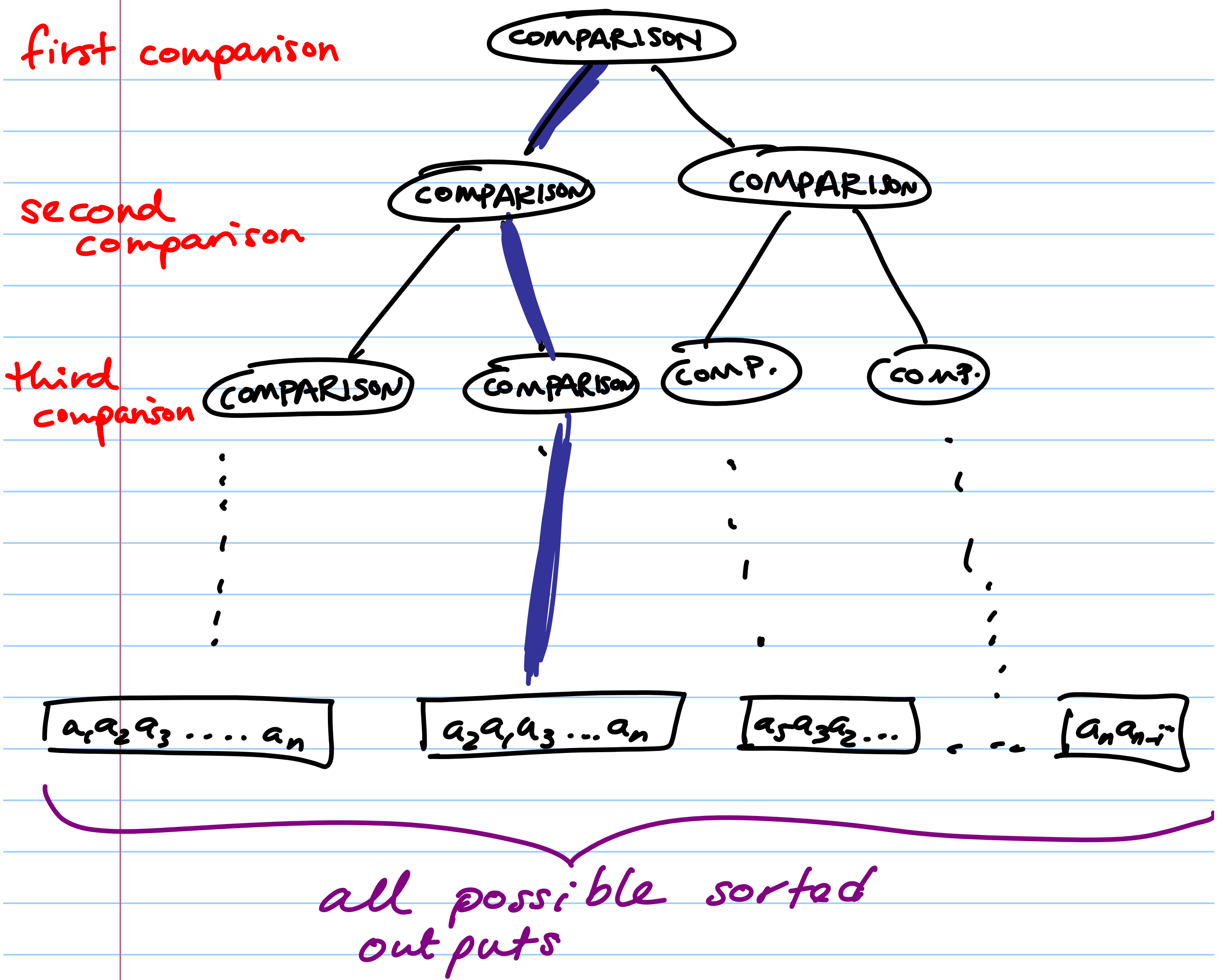
For Comparison-Based Sort

(1) Assume all input elements are distinct. (makes analysis easier)

(2) Then comparisons can be done with '<'.

(3) Model all comparison-based sorts as follows:

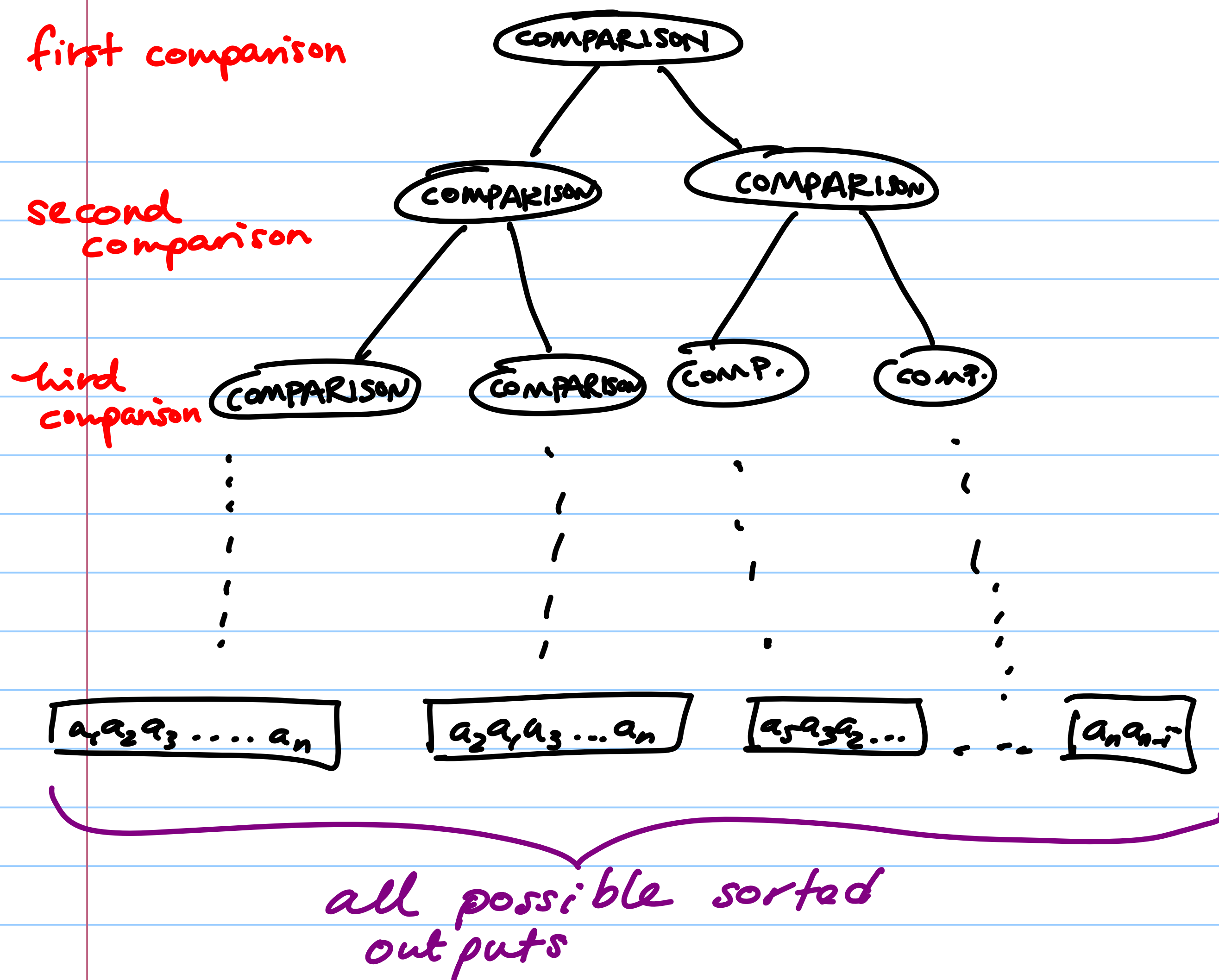
- input is $a_1\ a_2\ a_3 \ldots a_n$
- each node in tree is a comparison
- path through tree describes comparisons done in one execution of algorithm
- leaves are all possible outputs



if input is $a_1 a_2 a_3 a_4$ (say 4378)

output is $a_2 a_1 a_3 a_4$ (3478)

path is in blue



unique path to all outputs

How Many Outputs are there?

How Many Outputs are there?

Given input $a_1 a_2 \cdots a_n$

Possible outputs of a sort algorithm are:

how many?

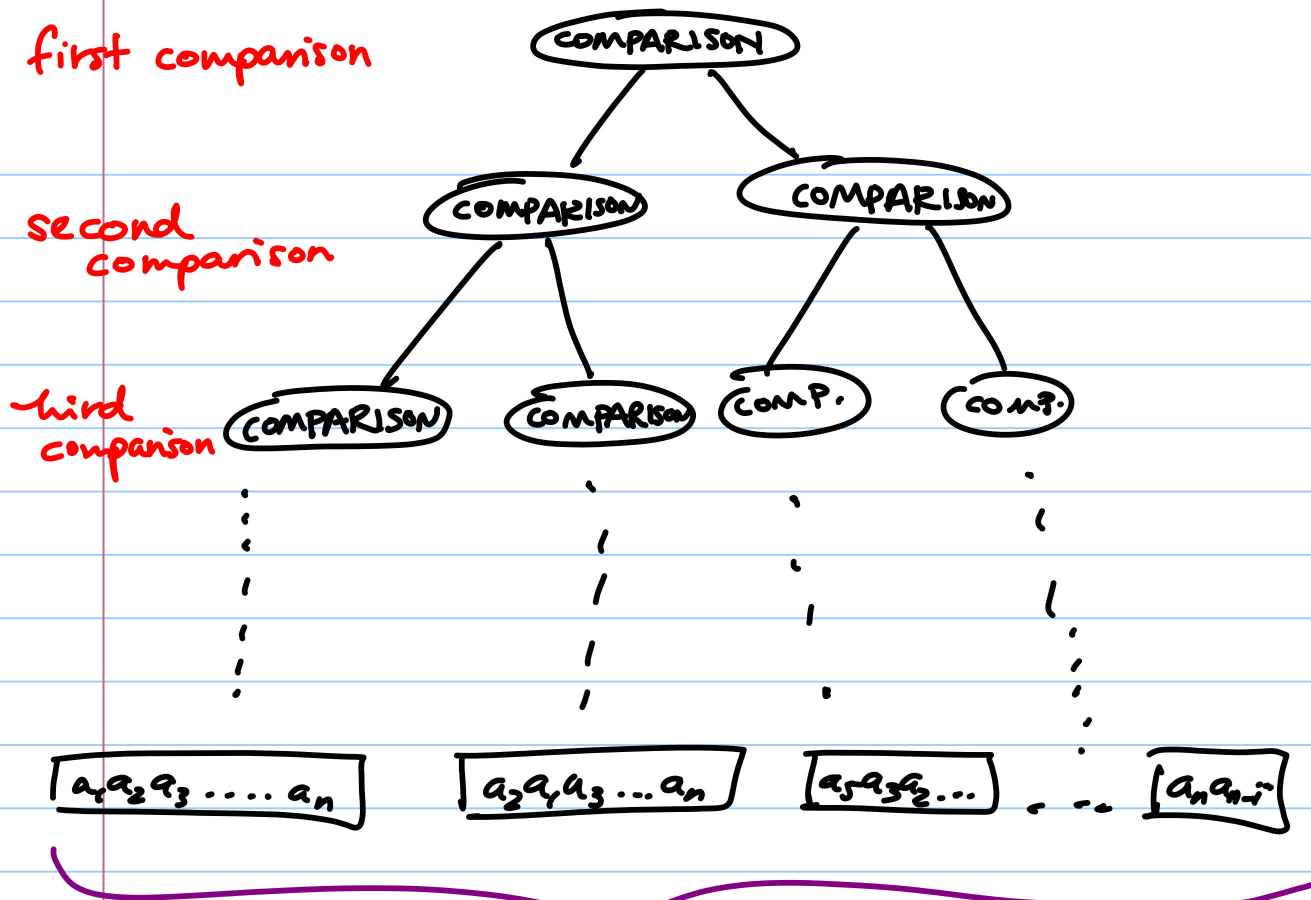
$$\left\{\begin{array}{l} a_1\ a_2\ \cdots\ a_n \\ a_2\ a_1\ \cdots\ a_n \\ a_3\ a_1\ \cdots\ n \\ \vdots \\ a_n\ a_{n-1}\ \cdots\ a_1 \end{array}\right.$$

$$n \cdot n-1 \cdot n-2 \cdot \cdots 1$$

possibilities for first value

possibilities for second value

$n!$ possible outputs



leaves = all possible sorted outputs

\# leaves $= n!$

max # comparisons in a path $= \log_2(n!)$
(#levels − 1)

Lower bound of sorting $\Theta(\log(n!))$

So in worst case, comparison based sorting inherently takes $\log_2(n!)$ comparisons!

How does that compare to the upper bound?

$$\log(n!) \in O(n \log n)$$

big O

$$\log(n!) \overset{?}{\leq} cn\log n$$

Note: $n! < n^n$

$n(n-1)(n-2)\cdots 1 < n \cdot n \cdot n \cdots n$

so

$$\log(n!) < \log(n^n)$$

new goal:

$$\log(n^n) \overset{?}{\leq} cn\log n$$

$$n \log n \leq cn\log n$$

Let $c = 1$, $n_0 = 1$.

$$\log(n!) \in \Omega(n \log n)$$

$$cn\log n \overset{?}{\leq} \log(n!)$$

$$cn\log n \overset{?}{\leq} \log\left[(n)(n-1)\dots 1\right]$$

$\log AB = \log A + \log B$

$$cn\log n \overset{?}{\leq} \log n + \log n-1 + \dots \underbrace{\log 1}_{\emptyset}$$

$$cn\log n \overset{?}{\leq} \underbrace{\sum_{i=1}^{n} \log i}_{\text{underestimate with integral}}$$

$u = \log_e i \quad dv = di$

$du = \frac{1}{i}\,di \quad v = i$

$$\int_1^n \log i \, di = i \log_e i \Big|_1^n - \int_1^n 1 \, di$$

$$= n \log_e n - n + 1$$

so new goal:

$$cn \log_e n \overset{?}{\leq} n \log_e n - n + 1$$

$$c n \log_e n \overset{?}{\leq} n \log_e n - n + 1$$

New goal:

$$c n \log_e n \overset{?}{\leq} n \log_e n - n$$

$$c \log_e n \leq \log_e n - 1$$

$$(c-1) \log_e n \leq -1$$

need $c < 1$

$$\underline{\text{let } c = \tfrac{1}{2}}$$

$$-\tfrac{1}{2} \log_e n \leq -1$$

$\log_e n = \dfrac{\log_2 n}{\log_2 e}$

$$\log_2 n \geq 2 \log_2 e$$

$$n \geq 2^{2 \log_2 e} = \left(2^{\log_2 e}\right)^2 = e^2 \approx 7.39$$

read 8.1 in book

$$\underline{\text{Let } n_0 = 8}$$

$$\boxed{\text{So} \quad \log(n!) \in \Theta(n \log n)}$$