COMP2054-ADE

Lecturer: Andrew Parkes
http://www.cs.nott.ac.uk/~pszajp/

Comparison Based Sorting

Lower Bound on Efficiency

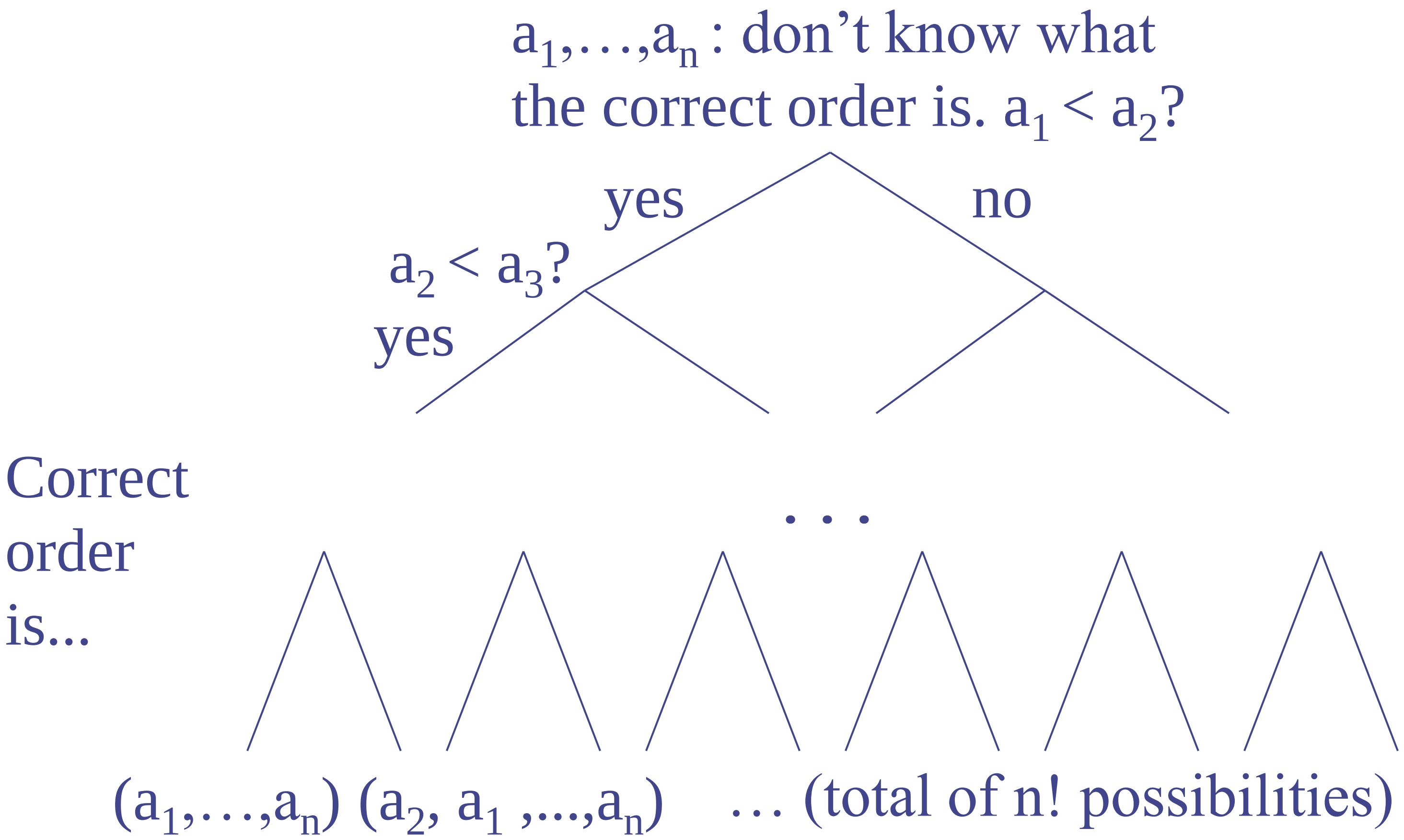
Comparison sorting

- A sorting algorithm is a comparison sorting algorithm if it uses comparisons between elements in the sequence to determine in which order to place them
- Examples of comparison sorts: bubble sort, selection sort, insertion sort, heap sort, merge sort, quicksort.
- Example of a sort that is ***not*** comparison-based: bucket sort
  - Runs in O(n), but relies on knowing the range of values in the sequence (e.g. “integers between 1 and 1000”).

Lower bound for comparison sort

- We can model sorting which depends on comparisons between elements as a binary decision tree.
- At each node, a comparison between two elements is made; there are two possible outcomes and we find out a bit more about the correct order of items in the array.
- Finally arrive at full information about the correct order of the items in the array.

Comparison sorting



How many comparisons?

- If a binary tree has n! leaves, than the minimal number of levels (assuming the tree is perfect) is $(\log_2 n!) + 1$.
- This shows that O(n log n) sorting algorithms are essentially optimal
  - $\log_2 n!$ is not equal to $n \log_2 n$, but has the same growth rate
- **Comparison-based sorting cannot do better than O(n log n)**
  - Technically, it uses Stirling's approximation https://en.wikipedia.org/wiki/Stirling%27s_approximation that

    $\log_2(n!) = n \log_2(n) +$ "smaller terms"

    - Note: you should know the this approximation, but (obviously) do not need to know the proof.

Questions to ask about sorting algorithms

- Big-Oh complexity (both time and space)?
  - Best case inputs? Worst case inputs?
- Extra workspace needed?
  Or is it `in-place'?
- Stable sorts?
- "Dynamic sorting" – how well does it do if the data is already "nearly sorted"
- Data access patterns?
  - Sequential? Random Access?
- Relevant and appropriate assertions

Aim to understand these issues for various sorting algorithms.

Minimum Expectations

- Know the meaning of ‘comparison-based sorting’ and the resulting O(n log n) lower bound on complexity.