Sorting Conclusions

David Kauchak
cs302
Spring 2012

Administrative

- Assignment 5 out today

Sorting bounds

- Mergsort is O(n log n)
- Quicksort is O(n log n) on average
- Can we do better?

Comparison-based sorting

- Sorted order is determined based **only** on a comparison between input elements
  - A[i] < A[j]
  - A[i] > A[j]
  - A[i] = A[j]
  - A[i] ≤ A[j]
  - A[i] ≥ A[j]
- Do any of the sorting algorithms we've looked at use additional information?

Comparison-based sorting

- Sorted order is determined based **only** on a comparison between input elements
  - A[i] < A[j]
  - A[i] > A[j]
  - A[i] = A[j]
  - A[i] ≤ A[j]
  - A[i] ≥ A[j]
- Do any of the sorting algorithms we've looked at use additional information?
  - No
  - All the algorithms we've seen are comparison-based sorting algorithms

Comparison-based sorting

- Sorted order is determined based **only** on a comparison between input elements
  - A[i] < A[j]
  - A[i] > A[j]
  - A[i] = A[j]
  - A[i] ≤ A[j]
  - A[i] ≥ A[j]
- In Java (and many languages) for a class of objects to be sorted we define a comparator
- What does it do?

Comparison-based sorting

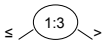
- Sorted order is determined based **only** on a comparison between input elements
  - A[i] < A[j]
  - A[i] > A[j]
  - A[i] = A[j]
  - A[i] ≤ A[j]
  - A[i] ≥ A[j]
- In Java (and many languages) for a class of objects to be sorted we define a comparator
- What does it do?
  - Just compares any two elements
  - Useful for comparison-based sorting algorithms

Comparison-based sorting

- Sorted order is determined based **only** on a comparison between input elements
  - A[i] < A[j]
  - A[i] > A[j]
  - A[i] = A[j]
  - A[i] ≤ A[j]
  - A[i] ≥ A[j]
- Can we do better than O(n log n) for comparison based sorting approaches?

Decision-tree model

- *Full* binary tree representing the comparisons between elements by a sorting algorithm
- Internal nodes contain indices to be compared



- Leaves contain a complete permutation of the input

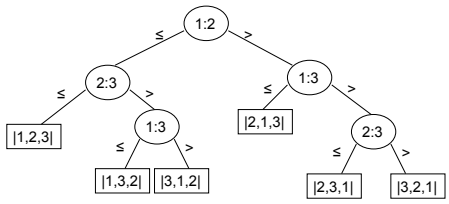
[3, 12, 7] ⇨ | 1,3,2 | ⇨ [3, 7, 12]

[7, 3, 12] ⇨ | 2,1,3 | ⇨ [3, 7, 12]

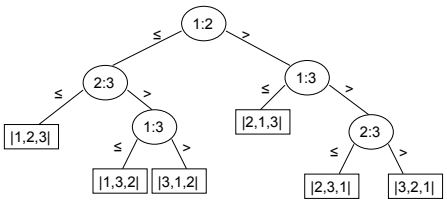
| 1,3,2 |

- Tracing a path from root to leave gives the correct reordering/permutation of the input for an input

A decision tree model

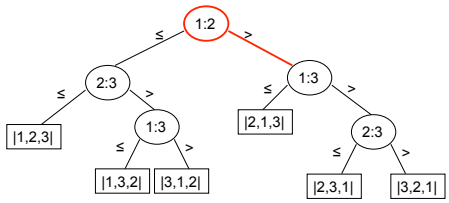


A decision tree model



[12, 7, 3]

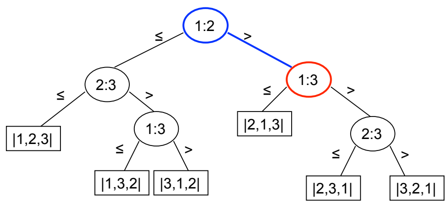
A decision tree model



[12, 7, 3]

Is 12 ≤ 7 or is 12 > 7?

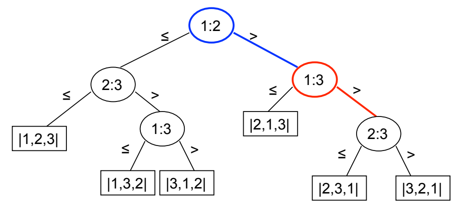
A decision tree model



[12, 7, 3]

Is 12 ≤ 3 or is 12 > 3?

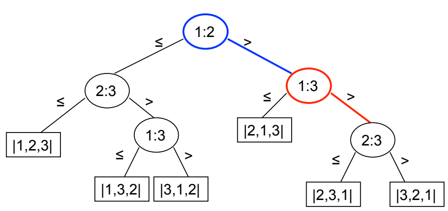
A decision tree model



[12, 7, 3]

Is 12 ≤ 3 or is 12 > 3?

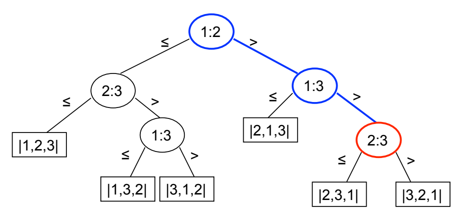
A decision tree model



[12, 7, 3]

Is 12 ≤ 3 or is 12 > 3?

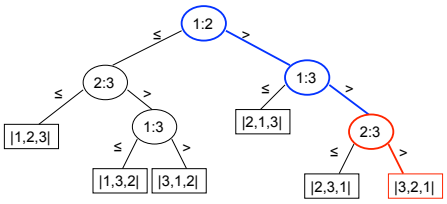
A decision tree model



[12, 7, 3]

Is 7 ≤ 3 or is 7 > 3?

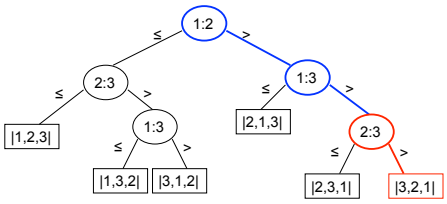
A decision tree model



[12, 7, 3]

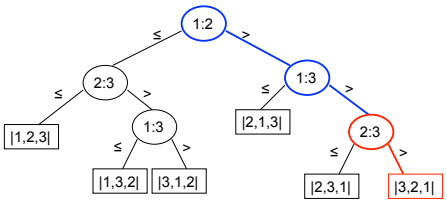
Is 7 ≤ 3 or is 7 > 3?

A decision tree model



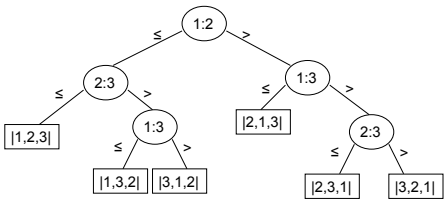
3, 2, 1
[12, 7, 3]

A decision tree model



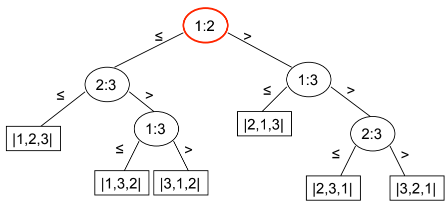
3, 2, 1
[12, 7, 3] ⇨ [3, 7, 12]

A decision tree model



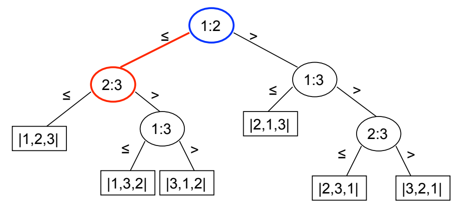
[7, 12, 3]

A decision tree model



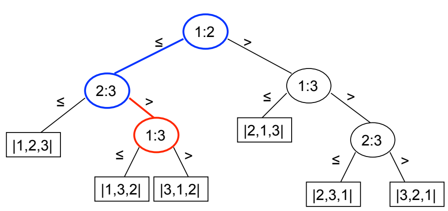
[7, 12, 3]

A decision tree model



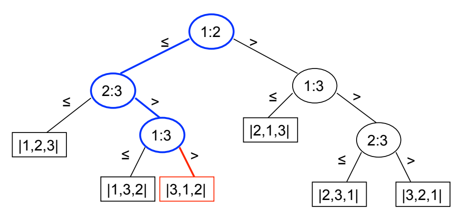
[7, 12, 3]

A decision tree model



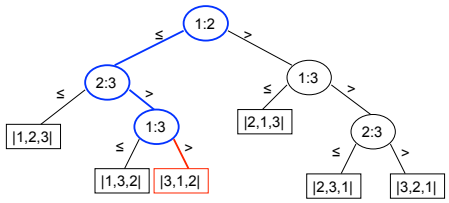
[7, 12, 3]

A decision tree model



[7, 12, 3]

A decision tree model



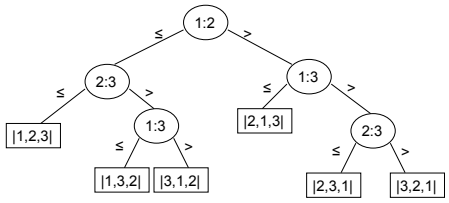
[7, 12, 3] ⇨ [3, 7, 12]

How many leaves are in a decision tree?

- Leaves **must** have all possible permutations of the input
- What if decision tree model didn't?
- Input of size *n*, *n!* leaves
- Some input would exist that didn't have a correct reordering

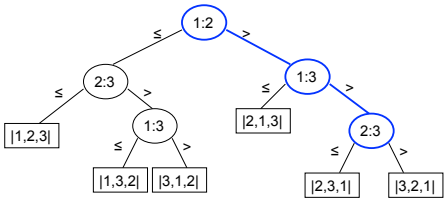
A lower bound

- What is the worst-case number of comparisons for a tree?



A lower bound

- The longest path in the tree, i.e. the height



A lower bound

- What is the maximum number of leaves a binary tree of height *h* can have?
- A complete binary tree has 2^h leaves

$$2^h \geq n!$$

$$h \geq \log n!$$ log is monotonically increasing

$$h = \Omega(n \log n)$$ from hw ☺

Can we do better?