Algorithms and Data Structures Week 6

Georg Hasebe, 25th October.

Binary Decision Trees

Let B be the binary decision tree of a comparison-based sorting algorithm $\mathcal{A}$.

Suppose we call $\mathcal{A}$ on an array with n numbers. For $\mathcal{A}$ to be a correct comparison-based sorting algorithm, it's binary decision tree must work on all possible permutations of these n numbers.

For example, consider the array $\{1, 2, 3\}$, then $\mathcal{A}$ must work on all six permutations (orderings):

$$(1,2,3),(1,3,2),(3,1,2),(2,1,3),(2,3,1),(3,2,1)$$

For an arbitrary array of length n, there are $n!$ of these permutations. In other words, our binary decision tree B must have $n!$ leaves (see figure 2.5). What is the height of this tree? It is $\log_2(n!)$, which is in $\Theta(n \log n)$ as you showed in exercise 2.5. Therefore, the height of any binary decision tree (no matter what comparison-based sorting algorithm we use), must be at least $n \log n$, i.e. in $\Omega(n \log n)$.

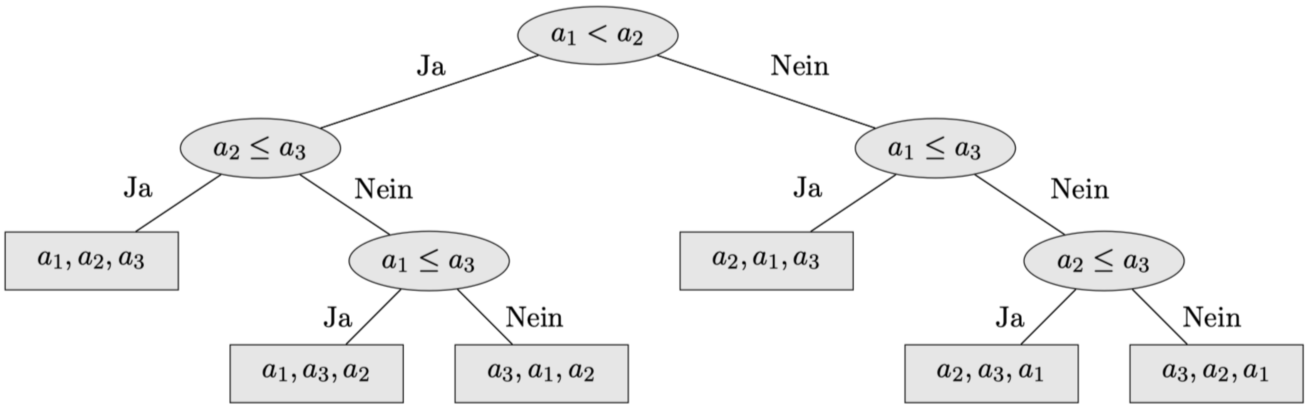

Abb. 2.5 Beispiel Entscheidungsbaum mit Array $[a_1, a_2, a_3]$

Consider the sorting-based algorithm *Bubble-Sort*. As you have seen in the lecture, the number of comparisons performed in this algorithm are in $O(n^2)$ and not $O(n \log n)$. In other words, *Bubble-Sort* is not able to figure out the ordering of some array using $O(n \log n)$ operations, but needs more, namely $O(n^2)$.

Coming back to today's quiz question: Let B be the binary decision tree of a comparison-based sorting algorithm for arrays of length n. Which of the following statements are true for B?

1. B contains at most 2^n nodes.

2. B contains at least $n!$ nodes.
3. B has depth at least $\Omega(n \log n)$.
4. B has depth at most $O(n \log n)$.

As we have shown, only 2. and 3. are true: for 1., we showed that the number of leaves must be $n! > 2^n$ for $n \geq 4$ (recall the leafs are only the nodes that don't have any other children, so the total amount of nodes is even more) and for 4. we showed that there exist comparison-based algorithms like *Bubble-Sort* that need more than $O(n \log n)$ comparisons.

Indeterminate Forms

Some of you have argued that, for a continuous function f we have:

$$\lim_{n\to\infty} f(n)^{g(n)} = f(n)^{\lim_{n\to\infty} g(n)}.$$

But that is not the case! In fact we have:

$$\lim_{n\to\infty} f(n)^{g(n)} = \lim_{n\to\infty} f(n)^{\lim_{n\to\infty} g(n)}.$$

(Notice the limit is both in the exponent and in front of the base).

Where this doesn't matter:

For example, for $\lim_{n\to\infty} e^{g(n)}$ we get $e^{\lim_{n\to\infty} g(n)}$ since $\lim_{n\to\infty} e$ is just e.

Where it does matter:

For example: $\lim_{n\to\infty} n^{\frac{3}{\ln n}}$. Clearly, we have $\lim_{n\to\infty} \frac{3}{\ln n} = 0$ and $n^0 = 1$ for all integers n, so one might argue that the entire limit evaluates to 1. This is **wrong**, since we'd have $\lim_{n\to\infty} n^0 = \infty^0$ and we don't really know what that is, do we? We don't know, because we don't know, how fast the exponent tends to 0 as n becomes really large. In other words, while the exponent becomes 0, the base is constantly growing and tends to infinity.

We use other strategies such as: $n^{\frac{3}{\ln n}} = e^{\frac{3}{\ln n} \cdot \ln n} = e^3$. It's now easy to see that $\lim_{n\to\infty} n^{\frac{3}{\ln n}} = e^3$.

What if we have a limit were we get something like $\infty \cdot 0$?

Some of you argued that it might be 0, but then we can prove $n^2 \leq O(n)$:

$$\lim_{n\to\infty} \frac{n^2}{n} = \lim_{n\to\infty} n^2 \cdot \frac{1}{n} = \infty \cdot 0 = 0.$$

Then $\infty \cdot 0$ must be ∞, right? But then:

$$\lim_{n\to\infty} \frac{n}{n^2} = \lim_{n\to\infty} n \cdot \frac{1}{n^2} = \infty \cdot 0 = \infty.$$

which is of course wrong.

This tells us that $\infty \cdot 0$ can mean multiple things, and we have to proceed differently.

Both of these examples lead to so-called indeterminate forms. Other examples include:

$$\frac{0}{0}, \frac{\infty}{\infty}, 0 \cdot \infty, \infty - \infty, \infty^0, 1^\infty, 0^0.$$

Note that these expressions are informal as in, they don't exist and don't mean anything. Read more about them on Wikipedia.