CSC165
O

GARY BAUMGARTNER

We'd like to be able to say that binary search is in some sense faster than linear search. But the time taken depends on exactly how we code the algorithms. And even if we decide on a particular implementation in code, we need to know how long the different instructions take. The effect of these details is roughly to multiply the time by some constant.

If we pick time units so that our binary search of n elements takes roughly $t_b(n) = \log n$ units of time, then linear search takes roughly $t_l(n) = cn$ units of time for some $c > 0$. We know from Calculus that t_b is eventually below t_l. So a binary search is going to be better than linear search for large enough numbers of elements, regardless of the exact implementations. Since it's usually only with large numbers of elements that the speed matters, this is a useful result.

What we make precise now is, for a function f:

The set of functions that are eventually no worse than a constant multiple of f.

To describe running times, we usual have an input size from N, and a positive time in R. We might want to talk about space as well, and then 0 becomes a possible output as well. So we restrict ourselves to functions from $N \to R^{\geq 0}$. Now we can give a precise definition.

For $f : N \to R^{\geq 0}$, let
$$O(f) = \left\{ g : N \to R^{\geq 0} | \exists c \in R^+, \exists b \in N, \forall n \in N, n \geq b \to g(n) \leq cf(n) \right\}.$$

"$O(f)$" is pronounced "big oh of f". We think of $O(f)$ informally as the functions that grow no faster than f.

We can summarize part of our discussion of binary and linear search by: $\log n \in O(n)$. CSC108 and CSC148 give you some more intuition and concrete examples. Here we do the precise details.

Let's prove that $3n^2 + 2 \in O(n^2)$. We need to find a point at and after which $3n^2 + 2$ is no more than some (positive) constant multiple of n^2. From $n = 2$ onward, $2 \leq n^2$ and so $3n^2 + 2 \leq 4n^2$. Let's now do this formally:

Let $c = 4$. Then $c \in R^+$.
Let $b = 2$. Then $b \in N$.
     Let $n \in N$.
     Suppose $n \geq b$.
         Then $n \geq b = 2$, so $n^2 \geq 2$, so $3n^2 + 2 \leq 3n^2 + n^2 = 4n^2 = cn^2$.
     Thus $n \geq b \to 3n^2 + 2 \leq cn^2$.
     Since $n \in N$ is arbitrary: $\forall n \in N, n \geq b \to 3n^2 + 2 \leq cn^2$.
Since $c \in R^+$ and $b \in N$:
     $\exists c \in R^+, \exists b \in N, \forall n \in N, n \geq b \to 3n^2 + 2 \leq cn^2$.
Therefore, $3n^2 + 2 \in O\left(n^2\right)$.

For another example, is $n^4 \in O\left(3n^2\right)$? Is n^4 eventually no more than some constant multiple of $3n^2$? No. No matter how big a constant c we use, n^4 is eventually bigger than $3cn^2$. More precisely:

$$(*) \; \forall c \in R^+, \forall b \in N, \exists n \in N, n \geq b \wedge n^4 > c \cdot 3n^2.$$

Let's do some rough work to determine what n we should pick based on c and b.

     We *want* $n^4 > 3cn^2$ for some $n \in N$ with $n \geq b$.
     This would be true if $n^2 > 3c$, $n^2 > 0$, $n \in N$ and $n \geq b$.
     This would be true if $n > \sqrt{3c}$, $n > 0$, $n \in N$ and $n \geq b$.
     This would be true if $n > \sqrt{3c}$, $n \geq b$ and $n \in N$, since $c > 0$.
     This would be true if $n = \left\lceil \max\left(1 + \sqrt{3c}, b\right) \right\rceil$.

Now we can do our proof that $n^4 \notin O\left(3n^2\right)$, by proving the negation $(*)$ (the negation of $n^4 \in O\left(3n^2\right)$):

Let $c \in R^+$.
Let $b \in N$.
     Let $n = \left\lceil \max\left(b, 1 + \sqrt{3c}\right) \right\rceil$.
     Since $\max\left(b, 1 + \sqrt{3c}\right) \geq 1 + \sqrt{3c} \geq 1$, it's ceiling n is in N.
         Let $m = \max\left(b, 1 + \sqrt{3c}\right)$.
         Then $m \geq b$, $m \geq 1 + \sqrt{3c} > 0$, and $m \geq 1 + \sqrt{3c} > \sqrt{3c}$.
         Since $n = \lceil m \rceil \geq m$, we have $n \geq b$, $n > 0$ and $n \geq \sqrt{3c}$.
         So $n^4 = n^2 n^2 > \left(\sqrt{3c}\right)^2 n^2$ (since $n \geq \sqrt{3c}$ and $n^2 > 0$) $= 3cn^2 = c \cdot 3n^2$.
     Thus $n \geq b \wedge n^4 > c \cdot 3n^2$.
     Since $n \in N$: $\exists n \in N, n \geq b \wedge n^4 > c \cdot 3n^2$.
     Since $m \geq b \geq 0$, it's ceiling is in N, so $n \in N$.
Since $c \in R^+$ and $b \in N$ are arbitrary:
     $\forall c \in R^+, \forall b \in N, \exists n \in N, n \geq b \wedge n^4 > c \cdot 3n^2$.
Therefore $n^4 \notin O\left(3n^2\right)$.

Can $f \in O(g)$ and $g \in O(f)$? Yes. We proved that $3n^2 + 2 \in O(n^2)$. It's also true (exercise!) that $n^2 \in O(3n^2 + 2)$. Even simpler, just let $f = g$.

Theorem. *If $f : N \to R^{\geq 0}$ then $f \in O(f)$.*

We said that $\log n \in O(n)$. We can also show that $n \in O(n^2)$. Thinking informally about time or space, we can say that $\log n$ is no worse than n, and n is no worse than n^2. This should let us conclude that $\log n$ is no worse than n^2. This is correct reasoning in general.

Theorem. *If $f, g, h : N \to R^{\geq 0}$, $f \in O(g)$ and $g \in O(h)$, then $f \in O(h)$.*

Before we prove this, let's play around. Suppose $f \in O(g)$ because for all $n \geq 200$ we have $f(n) \leq 3g(n)$, and suppose $g \in O(h)$ because for all $n \geq 400$ we have $g(n) \leq 5g(n)$. Then for all $n \geq 400$, $f(n) \leq 15h(n)$. In general, we get our b as the maximum of the two bs, and our c from the product of the two cs. Now we can do our proof.

> Suppose $f, g, h : N \to R^{\geq 0}$, $f \in O(g)$ and $g \in O(h)$.
> Then $f \in O(g)$, so $\exists c \in R^+, \exists b \in N, \forall n \in N, n \geq b \to f(n) \leq cg(n)$.
> Let $c_f \in R^+$ and $b_f \in N$ such that $\forall n \in N, n \geq b_f \to f(n) \leq c_f g(n)$.
> Also, $g \in O(h)$, so $\exists c \in R^+, \exists b \in N, \forall n \in N, g(n) \leq ch(n)$.
> Let $c_g \in R^+$ and $b_g \in N$ such that $\forall n \in N, n \geq b_g \to g(n) \leq c_g h(n)$.
> Let $c = c_f \cdot c_g$. Then $c \in R^+$ since $c_f, c_g \in R^+$.
> Let $b = \max(b_f, b_g)$. Then $b \in N$ since $b_f, b_g \in N$.
>> Let $n \in N$.
>> Suppose $n \geq b$.
>>> Then $n \geq b_f$, so $f(n) \leq c_f g(n)$.
>>> Also $n \geq b_g$, so $g(n) \leq c_g h(n)$.
>>> Thus $f(n) \leq c_f g(n) \leq c_f c_g h(n)$ (using $c_f \geq 0$ and the previous line) $= ch(n)$.
>> Thus $n \geq b \to f(n) \leq ch(n)$.
>> Since $n \in N$ is arbitrary: $\forall n \in N, n \geq b \to f(n) \leq ch(n)$.
> Since $c \in R^+$ and $b \in N$:
>> $\exists c \in R^+, \exists b \in N, \forall n \in N, n \geq b \to f(n) \leq ch(n)$.
> Thus $f \in O(h)$.
> Therefore $\left(f, g, h : N \to R^{\geq 0} \wedge f \in O(g) \wedge g \in O(h) \right) \to f \in O(h)$.