Big O Notation

July 21, 2014

Comparing functions

- Suppose the running time of program A is $f(n)$ and the running time of program B is $g(n)$, where n is the size of the input.
- How do we compare $f(n)$ and $g(n)$?
- Example: $f(n) = n$, $g(n) = n^2/100$.

- We only care about how the function grows, i.e., what happens to the functions as $n \to \infty$.
- If the function is a sum, we only care about the fastest growing term.

- $f \sim g$ if
$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$$
- $f \ll g$ if
$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$$

Example: $n^3 + 10 \sim n^3$ and $n^2 \ll n^3$.
Example: n^2 and $3n^2$ aren't comparable using this notation.
Example: $\sin(n)$ and 1 aren't comparable using this notation.

Ordering primitive functions

$$1 \ll \log n \ll n \ll n \log n \ll n^2 \ll \cdots \ll 2^n \ll 3^n \ll \cdots \ll n!$$

Proposition

For any positive integer $n \geq 4$, $\frac{2^n}{n!} < (1/2)^{n-4}$.

Dominant terms and constants

- When using $\ll$ and $\sim$, only the fastest-growing term matters
- When using $\ll$, multiplicative constants do not matter.
- Example: $n^3 + 2n^2 + 1000000 \sim n^3$ and $17n^2 \ll n^3$

Big O

- Sometimes we don't care about multiplicative constants.
- If f is non-negative, then $f(n)$ is $O(g(n))$ if there is a constant c such that $f(n) \leq cg(n)$ for large n.
- Example: $n^2 + n$ is $O(n^3)$ and $5n^3$ is $O(n^3)$.

Running time

When analyzing the running time of an algorithm, we want to find out how many "basic operations" are required, in terms of the input size *n*.

Example: Smallest distance

bestdist = infinity

for i = 1, ..., n

for j = 1, ..., n

if i != j

d = dist[i,j]

if bestdist ¿ d

bestdist

= d

A better algorithm?

for i = 1 to n for j = i+1 to n d = D[i,j] if bestdist ¿ d bestdist = d

Merging two lists

while L1 not empty or L2 not empty if head(L1) ¡= head(L2) or L2 is empty add head(L1) to end of L else add head(L2) to end of L if s = t return yes mark S and add it to while M is non-empty find some $p \in M$ for every neighbor q of p if $q = t$ return yes if q is unmarked mark q and add it to M

Square root

search area is from 1 to n guess middle if guess is too low, eliminate bottom half of search area if guess is too high, eliminate top half of search area.

Tower of hanoi

move top $n-1$ disks to pole 2 move big disk to pole 3 move top $n-1$ disks to pole 3.

Multiplication

The standard procedure for multiplying two n-digit numbers is $O(n^2)$.
Suppose $x = x_1 10^m + x_0$ and $y = y_1 10^m + y_0$, where x_1, x_0, y_1, y_0 are $(n/2)$-digit numbers.
Calculating

$$x_1 y p_1 10^{2m} + (x_0 y_1 + x_1 y_0) 10^m + x_0 y_0$$

requires 4 multiplications of $(n/2)$-digit numbers.

Karatsuba's algorithm

Instead, calculate $x_1 y_1$, $x_0 y_0$, and

$$(x_1 + x_0)(y_1 + y_0) - x_1 y_1 - x_0 y_0.$$

Calculating

$$x_1 yp_1 10^{2m} + (x_0 y_1 + x_1 y_0) 10^m + x_0 y_0$$

requires 3 multiplications of $(n/2)$-digit numbers.

NP

Problems in NP are yes/no problems where it is easy to verify a proof that the answer is "yes." However, these proofs may be hard to find in the first place.

Example: Is a graph 3-colorable?

NP-complete problems are the hardest problems in NP.

Proof by contradiction

- I want to prove that P is true.
- I assume that $\neg P$ is true.
- I prove that both Q and $\neg Q$ are true.
- This is impossible, so my initial assumption of $\neg P$ must be wrong.
- Thus, P is true.

$\sqrt{2}$

Proposition

$\sqrt{2}$ is irrational.

Infinitude of primes

Proposition

There are infinitely many prime numbers.

Compression

A compression algorithm is lossless if any input file can be reconstructed from its input file.

Proposition

There is no lossless compression algorithm that shrinks all files.

Sets of sets

Sets can contain sets as elements
Example: $\{\{1, 2\}, 3, \{3\}\}$

Power Sets

The power set of A is the set of all subsets of A and is denoted $\mathbb{P}(A)$.
For example, $\mathbb{P}(\{1, 2\}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$.

Russell's Paradox

Some sets can contain themselves.

For example, the set of all non-squares is not a square, so it should contain itself.

On the other hand, the set of all squares is not a square, so it does not contain itself.

Let S be the set of all sets that do not contain themselves. Does S contain itself?

Combinations

How many possible 5-card hands are there in a deck of 52 distinct cards?

We define

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Combinations

- How many 10-character strings of 0s and 1s have exactly 5 0s?
- How many 10-character strings of uppercase letters contain no more than 3 A's?
- How many ways are there to walk from the origin to (n, k) while only following the grid lines?

Balls and urns

I have 10 balls and 5 urns. The balls are indistinguishable but the urns are distinguishable. How many ways can I put the balls into the urns?

Pascal's triangle and identities

- $$\binom{n}{k} = \binom{n}{n-k}$$
- (Pascal's identity)
$$\binom{n-1}{k} + \binom{n-1}{k-1} = \binom{}{}$$
- (Hockey Stick identity)

Binomial theorem

What is $(x + y)^n$?