The Growth of Functions and Big-O Notation

Big-O Notation

Big-O notation allows us to describe the aymptotic growth of a function without concern for i) constant multiplicative factors, and ii) lower-order additive terms. For example, using big-O notation, the function $f(n) = 3n^2+6n+7$ is assumed to have the same kind of (quadratic) growth as $g(n) = n^2$.

Why do we choose to ignore constant factors and lower-order additive terms? One kind of function that we often consider throughout computing is $T(n)$, which represents the worst-case number of steps required by a an algorithm to process an input of size n. Function $T(n)$ will vary depending on the computing paradigm that is used to represent the algorithm. For example, one paradigm might represent the algorithm as a C program, while another might represent it as a sequence of random-access machine instructions. Now if T_1 measures the number of algorithmic steps for the first paradigm, and $T_2(n)$ measures the number of steps for the second, then, assuming that a paradigm does not include any unnecessary overhead, these two functions will likely be within multiplicative constant factors of one another. In other words, there will exist two constants C_1 and C_2 for which

$$C_1T_2(n) \leq T_1(n) \leq C_2T_2(n).$$

For this reason, big-O notation allows one to describe the steps of an algorithm in a mostly paradigm-indepenent manner, yet still be able to give meaningful representations of $T(n)$ by ignoring the paradigm-dependent constant factors.

Let $f(n)$ and $g(n)$ be functions from the set of nonnegative integers to the set of nonnegative real numbers. Then

Big-O $f(n) = \mathrm{O}(g(n))$ iff there exist constants $C > 0$ and $k \geq 1$ such that $f(n) \leq Cg(n)$ for every $n \geq k$.

Big-Ω $f(n) = \Omega(g(n))$ iff there exist constants $C > 0$ and $k \geq 1$ such that $f(n) \geq Cg(n)$ for every $n \geq k$.

Big-Θ $f(n) = \Theta(g(n))$ iff $f(n) = \mathrm{O}(g(n))$ and $f(n) = \Omega(g(n))$.

little-o $f(n) = \mathrm{o}(g(n))$ iff $\lim\limits_{n\to\infty} \frac{f(n)}{g(n)} = 0$.

little-ω $f(n) = \omega(g(n))$ iff $\lim\limits_{n\to\infty} \frac{f(n)}{g(n)} = \infty$.

Note that a more succinct way of saying "property $P(n)$ is true for all $n \geq k$, for some constant k" is to say "property $P(n)$ holds for *sufficiently large* n". Although this phrase will be used often throughout the course, nevertheless, when establishing a big-O relationship between two functions, the student should make the effort to provide the value of k for which the inequality is true.

Given functions $f(n)$ and $g(n)$, to determine the **big-O relationship** between f and g, we mean establishing which, if any, of the above growth relationships apply to f and g. Note that, if more than one of the above relations is true, then we choose the one that gives the most information. For example, if $f(n) = \mathrm{o}(g(n))$ and $f(n) = \mathrm{O}(g(n))$, then we would simply write $f(n) = \mathrm{o}(g(n))$, since it implies the latter relation.

Example 1. Determine the big-O relationship between i) $f(n) = 6n^2 + 2n + 5$ and $g(n) = 50n^2$, and ii) the same $f(n)$ and $g(n) = n^3$.

cubic

quadratic growth

Basic Results

In this section we provide some basic results that allow one to determine the big-O growth of a function, or the big-O relationship between two functions, without having to revert to the definitions.

Theorem 1. Let $p(n)$ be a polynomial of degree a and $q(n)$ be a polynomial of degree b. Then

- $p(n) = \mathrm{O}(q(n))$ if and only if $a \leq b$
- $p(n) = \Omega(q(n))$ if and only if $a \geq b$
- $p(n) = \Theta(q(n))$ if and only if $a = b$
- $p(n) = \mathrm{o}(q(n))$ if and only if $a < b$
- $p(n) = \omega(q(n))$ if and only if $a > b$

Thus, Theorem 1 could have been invoked to prove that $f(n) = \mathrm{o}(g(n))$, where f and g are the functions from Example 1.

Theorem 2. Let $f(n)$, $g(n)$, $h(n)$, and $k(n)$ be nonnegative integer functions for sufficiently large n. Then

- $f(n) + g(n) = \Theta(\max(f, g)(n))$
- if $f(n) = \Theta(h(n))$ and $g(n) = \Theta(k(n))$, then $f(n)g(n) = \Theta((hk)(n))$
- **Transitivity.** Let $R \in \{\mathrm{O}, \mathrm{o}, \Theta, \Omega, \omega\}$ be one of the five big-O relationships. Then if $f(n) = R(g(n))$, and $g(n) = R(h(n))$ then $f(n) = R(h(n))$. In other words, all five of the big-O relationships are transitive.

Example 2. Use the results of Theorems 1 and 2 to give a succinct expression for the big-O growth of $f(n)g(n)$, where $f(n) = n\log(n^4+1) + n(\log n)^2$ and $g(n) = n^2 + 2n + 3$. Note: by "succinct" we mean that no constants or lower-order additive terms should appear in the answer.

$g(n) = n^2 + 2n + 3 = \Theta(n^2)$

$f(n) = \Theta(\max(n\log(n^4+1),\ n\log^2 n))$

$n\log n^4+1 \approx n\log n^4 = 4n\log n$

$4n\log n$ vs. $n\log n\log n$

$\therefore f(n) = \Theta(n\log^2 n)$

$\therefore$ By Th. 2

$(fg)(n) = \Theta(n\log^2 n \cdot n^2)$

$= \Theta(n^3\log^2 n)$

Polylog cubic

Theorem 3. If $\lim_{n\to\infty} \frac{f(n)}{g(n)} = C$, for some constant $C > 0$, then $f(n) = \Theta(g(n))$.

Proof of Theorem 3. Mathematically,

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = C$$

means that, for every $\epsilon > 0$, there exists $k \geq 0$, such that

$$\left|\frac{f(n)}{g(n)} - C\right| < \epsilon,$$

small #

$\varepsilon = 0.1$

0.05

for all $n \geq k$. In words, $\frac{f(n)}{g(n)}$ can be made arbitrarily close to C with increasing values of n. Removing the absolute-value yields

$$C - \epsilon < \frac{f(n)}{g(n)} < C + \epsilon,$$

which implies

$$(C - \epsilon)g(n) < f(n) < (C + \epsilon)g(n).$$

Since $C > 0$ and $\epsilon > 0$ are constants, the latter inequalities imply $f(n) = \Theta(g(n))$ so long as $C - \epsilon > 0$. Therefore, choosing $\epsilon = C/2$, the result is proven.

Example 3. Suppose $a > 1$ and $b < 0$ are constants, with $|b| < a$. Prove that $a^n + b^n = \Theta(a^n)$.

$$\lim_{n\to\infty} \frac{a^n + b^n}{a^n} = \lim_{n\to\infty} 1 + \frac{b^n}{a^n} = \lim_{n\to\infty} 1 + \left(\frac{b}{a}\right)^n = 1 + 0 = 1 > 0$$

$\therefore$ By Th. 3, $a^n + b^n = \Theta(a^n)$

exponential growth

L'Hospital's Rule. Suppose $f(n)$ and $g(n)$ are both differentiable functions with either

1. $\lim_{n\to\infty} f(n) = \lim_{n\to\infty} g(n) = \infty$, or
2. $\lim_{n\to\infty} f(n) = \lim_{n\to\infty} g(n) = 0$.

Then

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

$\log_e 2 = \ln 2$

Example 4. Prove that for every $\epsilon > 0$, $\log n = o(n^\epsilon)$. Note: in general, $\log^k n = o(n^\epsilon)$, where $k > 0$ is any integer constant.

Change of base

$$\log_a n = \frac{\log_b n}{\log_b a}$$

constant

$$\ln n = \log_e n$$

$$\lim_{n\to\infty} \frac{\ln n}{n^\varepsilon} = \lim_{n\to\infty} \frac{(\ln n)'}{(n^\varepsilon)'}$$

$$0$$

$$= \lim_{n\to\infty} \frac{\frac{1}{n}}{\varepsilon n^{\varepsilon-1}} = \frac{1}{\varepsilon} \lim_{n\to\infty} \frac{1}{n^\varepsilon} = 0$$

$$n^\varepsilon \to \infty$$

$$\therefore \log n = o(n^\varepsilon)$$

The following terminology is often used to describe the big-O growth of a function.

Growth	**Terminology**
$\Theta(1)$	constant growth
$\Theta(\log n)$	logarithmic growth
$\Theta(\log^k n)$, for some integer $k \geq 1$	polylogarithmic growth
$\Theta(n^k)$ for some positve $k < 1$	sublinear growth
$\Theta(n)$	linear growth
$\Theta(n \log n)$	log-linear growth
$\Theta(n \log^k n)$, for some integer $k \geq 1$	polylog-linear growth
$O(n^k)$ for some integer $k \geq 1$	polynomial growth
$\Omega(n^k)$, for every integer $k \geq 1$	superpolynomial growth
$\Omega(a^n)$ for some real $a > 1$	exponential growth

$T(n)$

$\downarrow$

running time function

Example 5. Use the above terminology to describe the growth of the functions from Examples 1 and 2.

Series and Summations

When analyzing a data structure or algorithm, quite often we will encounter a a **series**, which is an expression of the form

$$\sum_{i=1}^{n} f(i) = f(1) + f(2) + \cdots + f(n),$$

for some function $f(i)$. For example, if $f(i) = 2i$ and $n = 6$, then

$$\sum_{i=1}^{n} f(i) = \sum_{i=1}^{6} 2i = 2 + 4 + 6 + 8 + 10 + 12 = 42.$$

Note that f is called the **summand**, i the **index variable**, and n the **summation limit**.

In most applications the value of n is not given. Rather, we must determine the growth of the **sum function**

$$S(n) = \sum_{i=1}^{n} f(i)$$

that provides the sum of the series for a given positive integer n. For some series, the value of $S(n)$ can be given with a formula, such as the ones below.

Constant Sum $\sum_{i=1}^{n} 1 = n$

Arithmetic Sum $\sum_{i=1}^{n} i = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$.

Sum of Squares $\sum_{i=1}^{n} i^2 = 1 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$. $= \Theta(n^3)$

Sum of Cubes $\sum_{i=1}^{n} i^3 = 1 + 2^3 + \cdots + n^3 = [\frac{n(n+1)}{2}]^2$.

Geometic Series $\sum_{i=0}^{n} ar^i = a + ar + ar^2 + \cdots + ar^n = \frac{a(r^{n+1}-1)}{r-1}$.

Linear Combination

$$\sum_{i=1}^{n} (af(i) + bg(i)) = a\sum_{i=1}^{n} f(i) + b\sum_{i=1}^{n} g(i),$$

where a and b are constants, and f and g are functions that depend on i.

Example 6. Use the above formulas to evaluate the summation expression

$$\sum_{i=1}^{n} (7i^2 + i + 8). =$$

The final answer should be an expression that depends only on n.

$$7\sum_{i=1}^{n} i^2 + \sum_{i=1}^{n} i + 8\sum_{i=1}^{n} 1 =$$

$$\Theta(n^3) \qquad \Theta(n^2) \qquad \Theta(n)$$

$$7\,\frac{n(n+1)(2n+1)}{6} + \frac{n(n+1)}{2} + 8n$$

Exercise: Simplify as a polynomial $an^3 + bn^2 + cn + d$

Example 7. Evaluate the summation expression

$$\sum_{i=1}^{n}\sum_{j=1}^{i}(3i+2j).$$

The final answer should be an expression that depends only on n.

$$\sum_{j=1}^{i}(3i+2j) = 3i\sum_{j=1}^{i} 1 + 2\sum_{j=1}^{i} j =$$

$$3i^2 + 2\frac{i(i+1)}{2} = 3i^2 + i^2 + i = 4i^2 + i$$

$$\sum_{i=1}^{n}(4i^2+i) = 4\sum_{i=1}^{n} i^2 + \sum_{i=1}^{n} i =$$

$$\frac{4n(n+1)(2n+1)}{6} + \frac{n(n+1)}{2}$$

Unfortunately, the sum function $S(n)$ for many important series cannot be expressed using a formula. However, the next result shows how we may still determine the big-O growth of $S(n)$, which quite often is our main interest.

Integral Theorem. Let $f(x) > 0$ be an increasing or decreasing Riemann-integrable function over the interval $[1, \infty)$. Then

$$\sum_{i=1}^{n} f(i) = \Theta(\int_1^n f(x)dx),$$

$2^{2^{2^n}}$

if f is decreasing. Moreover, the same is true if f is increasing, provided $f(n) = O(\int_1^n f(x)dx)$.

Proof of Integral Theorem. We prove the case when f is decreasing. The case when f is increasing is left as an exercise. The quantity $\int_1^n f(x)dx$ represents the area under the curve of $f(x)$ from 1 to n. Moreover, for $i = 1, \ldots, n-1$, the rectangle R_i whose base is positioned from $x = i$ to $x = i+1$, and whose height is $f(i+1)$ lies under the graph. Therefore,

$$\sum_{i=1}^{n-1} \text{Area}(R_i) = \sum_{i=2}^{n} f(i) \le \int_1^n f(x)dx.$$

Adding $f(1)$ to both sides of the last inequality gives

$$\sum_{i=1}^{n} f(i) \le \int_1^n f(x)dx + f(1).$$

Now, choosing $C > 0$ so that $f(1) = C\int_1^n f(x)dx$ gives

$$\sum_{i=1}^{n} f(i) \le (1+C)\int_1^n f(x)dx,$$

which proves $\sum\limits_{i=1}^{n} f(i) = O(\int_1^n f(x)dx)$.

Now, for $i = 1, \ldots, n-1$, consider the rectangle R'_i whose base is positioned from $x = i$ to $x = i+1$, and whose height is $f(i)$. This rectangle covers all the area under the graph of f from $x = i$ to $x = i+1$. Therefore,

$$\sum_{i=1}^{n-1} \text{Area}(R'_i) = \sum_{i=1}^{n-1} f(i) \ge \int_1^n f(x)dx.$$

Now adding $f(n)$ to the left side of the last inequality gives

$$\sum_{i=1}^{n} f(i) \ge \int_1^n f(x)dx,$$

which proves $\sum\limits_{i=1}^{n} f(i) = \Omega(\int_1^n f(x)dx)$.

Therefore,

$$\sum_{i=1}^{n} f(i) = \Theta(\int_1^n f(x)dx).$$

Determine the big-O growth of the series Harmonic Series

$$\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n}.$$

$$f(i) = \frac{1}{i}$$

$$S(n) = \sum_{i=1}^{n} \frac{1}{i} = \Theta\left(\int_1^n \frac{1}{x}\,dx\right) = \Theta(\ln n)$$

$$\int_1^n \frac{1}{x}\,dx = \ln x \Big|_1^n = \ln n - \ln 1 = \ln n$$

$$\max(f,g) \le (f+g)(n) \le 2\max(f,g)$$

$$\Omega \qquad \qquad O$$

$$(f+g)(n) \ge 1 \max(f,g)(n)$$

$$= C$$

$$20 + 30 \le 2 \cdot 30$$

$$=$$

$$(f+g)(n) \le C \max(f,g)(n)$$

$$(f+g)(n) = O(\max(f,g)(n))$$

Exercises

1. Use the definition of big-Ω to prove that $n \log n = \Omega(n + n \log n^2)$. Provide appropriate C and k constants.

2. Provide the big-O relationship between $f(n) = n \log n$ and $g(n) = n + n \log n^2$.

3. Prove that $f(n) = O(g(n))$ if and only if $g(n) = \Omega(f(n))$.

4. Use the definition of big-Θ to prove that $f(n) + g(n) = \Theta(\max(f(n), g(n)))$.

5. Prove that $(n + a)^b = \Theta(n^b)$, for all real a and $b > 0$. Explain why Theorem 1 and L'Hospital's rule be avoided when solving this problem?

6. Prove that if $\text{limit}_{n\to\infty} \frac{f(n)}{g(n)} = 0$, then $f(n) = O(g(n))$, but $g(n) \neq O(f(n))$.

7. Prove or disprove: $2^{n+1} = O(2^n)$.

8. Prove or disprove: $2^{2n} = O(2^n)$.

9. Use any techniques or results from lecture to determine a succinct big-Θ expression for the growth of the function $\log^{50}(n)n^2 + \log(n^4)n^{2.1} + 1000n^2 + 100000000n$.

10. Prove or disprove: if $f(n) = O(g(n))$, then $2^{f(n)} = O(2^{g(n)})$.

11. Prove transitivity of big-O: if $f(n) = O(g(n))$, then $g(n) = O(h(n))$, then $f(n) = O(h(n))$.

12. If $g(n) = o(f(n))$, then prove that $f(n) + g(n) = \Theta(f(n))$.

13. Use L'Hospital's rule to prove Theorem 1. Hint: assume a and b are nonnegative integers and that $a \geq b$.

14. Use L'Hospital's rule to prove that $a^n = \omega(n^k)$, for every real $a > 1$ and integer $k \geq 1$.

15. Prove that $\log_a n = \Theta(\log_b n)$ for all $a, b > 0$.

16. Simplify each summation to an expression whose only variable is n, and provide the big-O growth of the expression.

    a. $\sum_{i=1}^{n}(n - 2i + 3)$
    b. $\sum_{i=0}^{n-1}(4i^2 - 2i + 7)$
    c. $\sum_{j=10}^{n} j$
    d. $\sum_{i=1}^{n}\sum_{j=1}^{n} j$
    e. $\sum_{i=1}^{n}\sum_{j=i}^{n}(j - i)$

17. Suppose $g(n) \geq 1$ for all n, and that $f(n) \leq g(n) + L$, for some constant $L \geq 0$ and all n. Prove that $f(n) = O(g(n))$.

18. Give an example that shows that the statement of Exercise 17 may not be true if we no longer assume $g(n) \geq 1$.

19. Use the Integral Theorem to establish that $1^k + 2^k + \cdots + n^k = \Theta(n^{k+1})$, where $k \geq 1$ is an integer constant.

20. Use the Integeral Theorem to prove that $\log 1 + \log 2 + \cdots + \log n = \Theta(n \log n)$.

21. Show that $\log(n!) = \Theta(n \log n)$.

22. Determine the big-O growth of $n + n/2 + n/3 + \cdots + 1$.

Exercise Hints and Solutions

1. Answers may vary. For this solution, $n + n\log n^2 \leq n\log n + 2n\log n = 3n\log n$. Thus, $n\log n \geq (1/3)(n + n\log n^2)$, for all $n \geq 1$. So $C = 1/3$ and $k = 1$.

2. From the previous exercise we have $f(n) = \Omega(g(n))$. But $f(n) = \mathrm{O}(g(n))$ with $C = k = 1$. Thus, $f(n) = \Theta(n)$.

3. Set up the inequality for big-O and divide both sides by C.

4. Two inequalities must be established, and $C_1 = 0.5$, $k_1 = 1$, $C_2 = 2$, $k_2 = 1$ are adequate constants (why?).

5. Use Theorem 3.

6. Since
$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0,$$
we know that $f(n) \leq g(n)$ for sufficiently large n. Thus, $f(n) = \mathrm{O}(g(n))$.

   Now suppose it was true that $g(n) \leq Cf(n)$ for some constant $C > 0$, and n sufficiently large. Then dividing both sides by $g(n)$ yields $\frac{f(n)}{g(n)} \geq 1/C$ for sufficiently large n. But since
$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0,$$
we know that $\frac{f(n)}{g(n)} < 1/C$, for sufficiently large n, which is a contradiction. Therefore, $g(n) \neq \mathrm{O}(f(n))$.

7. True, since $2^{n+1} = 2 \cdot 2^n$. $C = 2$, $k = 1$.

8. False. $2^{2n} = 4^n$ and
$$\lim_{n\to\infty} \frac{2^n}{4^n} = \lim_{n\to\infty} \frac{1}{2^n} = 0.$$
Now use Exercise 6.

9. $\Theta(n^{2.1}\log n)$.

10. False. Consider $f(n) = 2n$ and $g(n) = n$.

11. Assume $f(n) \leq C_1 g(n)$, for all $n \geq k_1$ and $g(n) \leq C_1 h(n)$, for all $n \geq k_2$. Therefore,
$$f(n) \leq C_1 g(n) \leq C_1 C_2 h(n)$$
for all $n \geq \max(k_1, k_2)$. $C = C_1 C_2$ and $k = \max(k_1, k_2)$.

12. Use Theorem 2 and the fact that $g(n) < f(n)$ for n sufficiently large.

13. Let $f(n)$ be a nonnegative-valued polynomial of degree a, and $g(n)$ be a nonnegative-valued polynomial of degree b, with $a \geq b$. Since the k th derivative (as a function of n) of $f(n)$ is equal to $\Theta(n^{a-k})$ and that of $g(n)$ is equal to $\Theta(n^{b-k})$ it follows that L'Hospital's rule applies to the ratio $f^{[k]}(n)/g^{[k]}(n)$ for all $k < b$. Moreover, upon applying the rule for the b th time, the numerator will equal a polyhomial of degree $a - b$, while the denominator will equal a constant. Thus $f(n) = \Theta(g(n))$ if $a = b$ and $f(n) = \omega(g(n))$ if $a > b$.

14. Since the derivative (as a function of n) of a^n equals $(\ln a)a^n$, it follows that the k th derivative of a^n divided by the k th derivative of n^k equals $\ln^k a^n/k!$, which tends to infinity. Therefore, $a^n = \omega(n^k)$.

15. By the Change of Base formula,

$$\log_a n = \frac{\log_b n}{\log_b a},$$

and so $\log_a n = C \log_b n$, where $C = 1/\log_b a$. Therefore, $\log_a n = \Theta(\log_b n)$.

16. Note: the final expressions have been simplified (which the exercise did not require).

   a. $\sum_{i=1}^{n}(n - 2i + 3) = n^2 - n(n+1) + 3n = 2n = \Theta(n)$.
   b. $\sum_{i=0}^{n-1}(4i^2 - 2i + 7) = \frac{2(n-1)(n)(2n-1)}{3} - n(n-1) + 7n = \Theta(n^3)$.
   c. $\sum_{j=10}^{n} j = \frac{n(n+1)}{2} - \frac{9(10)}{2} = \Theta(n^2)$.
   d. $\sum_{i=1}^{n}\sum_{j=1}^{n} j = \sum_{i=1}^{n}(n^2/2 + n/2) = n^3/2 + n^2/2 = \Theta(n^3)$.
   e. $\sum_{i=1}^{n}\sum_{j=i}^{n}(j - i) = \sum_{i=1}^{n}(n^2/2 + n/2 - i^2/2 + i/2 - ni + i^2 - i)$. Then evaluate the outer sum to get $\Theta(n^3)$.

17. $f(n) \le g(n) + L \le g(n) + Lg(n) \le (1 + L)g(n)$, for all $n \ge 0$, where the second inequality is true since $g(n) \ge 1$. $C = 1 + L$ and $k = 1$. The result still holds so long as $g(n)$ is bounded away from zero.

18. Consider $f(n) = 1/n$ and $g(n) = 1/n^2$.

19. By the Integral Theorem,

$$\sum_{i=1}^{n} i^k = \Theta(\int_1^n x^k dx) = \Theta(\frac{x^{k+1}}{k+1}\Big|_1^n) = \Theta(n^{k+1}).$$

20. By the Integral Theorem,

$$\sum_{i=1}^{n} \ln i = \Theta(\int_1^n \ln x dx).$$

Moreover,

$$\int_1^n \ln x dx = x \ln x|_1^n - \int_1^n 1 dx =$$
$$n \ln n - n + 1 = \Theta(n \ln n).$$

21. Since $\log ab = \log a + \log b$, we have

$$\log(n!) = \log(n(n-1)(n-2)\cdots 1) = \log n + \log(n-1) + \log(n-2) + \cdots \log 1.$$

Therefore, from the previous exercise, we have $\log(n!) = \Theta(n \log n)$.

22. We have

$$n + n/2 + n/3 + \cdots + 1 = n(1 + 1/2 + 1/3 + \cdots + 1/n).$$

Therefore, by Example 3, $n + n/2 + n/3 + \cdots + 1 = \Theta(n \log n)$.