> *"It's very illuminating to think about the fact that some — at most four hundred — years ago, professors at European universities would tell the brilliant students that if they were very diligent, it was not impossible to learn how to do long division. You see, the poor guys had to do it in Roman numerals. Now, here you see in a nutshell what a difference there is in a good and bad notation."*
>
> — Edsger W. Dijkstra
> *Datamation* Vol.23, No.5, p.164, 1977

> *"Make it as simple as possible. But no simpler."*
>
> — Albert Einstein
> (paraphrase)

# Lecture II
# RECURRENCES

Recurrences occur frequently in mathematics. In computer science, they arise through combinatorics, probability and analysis of algorithms. This chapter treats a class of recurrences arising from the analysis of recursive algorithms.
We begin with some working rules for solving recurrences, stressing the use of real recurrences and $\Theta$-order analysis. The latter emphasis leads to elementary (non-calculus) tools. We call this the *real elementary approach* to recurrences, recommended for its pedagogic advantages. The highlight of this chapter are two "Master Theorems".
We also have an Appendix to recall some basic properties of the exponential and logarithm functions.

## §1. Recurrences in Analysis of Algorithms

**¶1.** Although recurrences arise in many areas of mathematics, we shall largely look at those that arise from the complexity analysis of algorithms, and in combinatorial analysis. We introduce some basic techniques for solving such recurrences. A recurrence is a recursive relation for a complexity function $T(n)$. Here are two examples:

$$F(n) = F(n-1) + F(n-2) \tag{1}$$

and

$$T(n) = n + 2T(n/2). \tag{2}$$

Fibonacci in nature

The reader may recognize the first as the recurrence for Fibonacci numbers, and the second as the complexity of the Mergesort, described in Chapter 1, §I.6. These recurrences have[1] the following "separable form":

$$T(n) = G(n, T(n_1), \ldots, T(n_k)) \tag{3}$$

[1] Non-separable recurrences looks like $G(n, T(n), T(n_1), \ldots, T(n_k)) = 0$, but these are rare.

where $G(x_0, x_1, \ldots, x_k)$ is a function in $k+1$ variables and each $n_i = n_i(n)$ $(i = 1, \ldots, k)$ is a function of $n$ that is strictly less than $n$. E.g., in (1), we have $k = 2$ and $n_1 = n-1, n_2 = n-2$ while in (2), we have $k = 1$ and $n_1 = n/2$.

**¶2. Our approach to recurrences** What does it mean to "solve" recurrences such as equations (1)–(2)? The Fibonacci and Mergesort recurrences have the following well-known solutions:

$$F(n) = \Theta(\phi^n) \tag{4}$$

where $\phi = (1+\sqrt{5})/2 = 1.618\ldots$ is the golden ratio, and

*Solve up to Θ-order*

$$T(n) = \Theta(n \log n).$$

In this book, we generally estimate complexity functions $T(n)$ only[2] up to its Θ-order. The reason goes back to Chapter I, where we saw the importance of robustness properties in complexity results. If only an upper bound or lower bound is needed, then we determine $T(n)$ up to its $\mathcal{O}$-order or to Ω-order. In rare cases, we may be able to derive the exact solution (in fact, this is possible for $T(n)$ and $F(n)$ above). One benefit of Θ-order solutions is this — most of the recurrences we treat in this book can be solved by purely elementary methods (i.e., without use of calculus). This means that we do not have to assume differentiability of our functions, just Θ-bounds. We call our approach the **Real Elementary Approach** to recurrences. We just explained the "elementary" part of the approach, but what about the "real" part?

**¶3.** The variable "$n$" is called the **designated variable** of the recurrence (3). If there are non-designated variables, they are supposed to be held constant. In mathematics, we usually reserve "$n$" for natural numbers or perhaps integers. In the above examples, this is the natural interpretation for $n$. But one of the first steps we take in solving recurrences is to re-interpret $n$ (or whatever is the designated variable) to range over the real numbers. The corresponding recurrence equation (3) is then called a **real recurrence**. For this reason, we may prefer the symbol "$x$" as our designated variable, since $x$ is normally viewed as a real variable. This accounts for the 'real' part of our "Real Elementary Approach".

*get real!*

What does extending an integer recurrence $T(n)$ to a real recurrence $T(x)$ really produce? In the Fibonacci recurrence (1), what is $F(1.23)$? In Mergesort (2), what is the significance of $T(\pi) = T(3.14159\ldots)$? The short answer is, we don't really care. What is important is that the arguments $x$ for which we care about (e.g., $x \in \mathbb{N}$) are still faithfully captured by this extension.

In addition to the recurrence (3), we generally need the **boundary conditions** or **initial values** of the function $T(n)$. They give us the values of $T(n)$ *before* the recurrence (3) becomes valid. Without initial values, a recurrence is generally under-determined. For our example (1), if $n$ ranges over natural numbers, then the initial conditions

$$F(0) = 0, \qquad F(1) = 1$$

*Some initial conditions yield trivial solutions...*

give rise to the **standard Fibonacci numbers**, *i.e.*, $F(n)$ is the $n$th Fibonacci number where[3]

$$F(0,1,2,3,4,\ldots) = (0,1,1,2,3,\ldots) \tag{5}$$

In the literature, the initial conditions $F(0) = F(1) = 1$ is sometimes used, giving rise to a sequence that is just a shift-by-one of the standard Fibonacci numbers. On the other hand, if we use the initial conditions $F(0) = F(1) = 0$, then the solution is trivial: $F(n) = 0$ for all $n \geqslant 0$. Thus, our assertion earlier that $F(n) = \Theta(\phi^n)$ is the solution to (1) is not really true without knowing the initial conditions! On the other hand, our assertion $T(n) = \mathcal{O}(n \log n)$ (see Exercise) can be shown to hold for (2) *regardless of the initial conditions.*

What causes this difference between $T(n)$ and $F(n)$? It is that $T(n)$ is **non-homogeneous** while $F(n)$ is **homogeneous**. We can make $F(n)$ non-homogeneous by adding a non-zero term $f(n)$ such as $f(n) = 1$ or $f(n) = n$:

$$F(n) = F(n-1) + F(n-2) + f(n).$$

The resulting solution would again be independent of initial conditions. Indeed, most recurrences that arise in complexity analysis are non-homogeneous and they are essentially independent of initial conditions. *We shall assume and take advantage of this property* in our treatment of recurrences in this chapter.

## Exercises

**Exercise 1.1:** The standard Fibonacci sequence is $F(0,1,2,3,4,5,\ldots) = (0,1,1,2,3,5,\ldots)$. Give the recurrence for $G(n)$ so that $G(n) = -F(n)$ for all $n \geqslant 0$. E.g., $G(0,1,2,3,4,5) = (0,-1,-1,-2,-3,-5)$. ♢

**Exercise 1.2:** Why is the recurrence $G_0(n+1) = G_0(n) - G_0(n-1)$ not interesting? What about the non-homogeneous case, $G(n+1) = \delta + G(n) - G(n-1)$ where $\delta > 0$? ♢

**Exercise 1.3:** Consider the non-homogeneous version of Fibonacci recurrence $F(n) = F(n-1) + F(n-2) + f(n)$ for some function $f(n)$. If $f(n) = 1$, show that $F(n) = \Omega(c^n)$ for some $c > 1$, regardless of the initial conditions. Try to find the largest value for $c$. Does your bound hold if we have $f(n) = n$ instead? ♢

**Exercise 1.4:** Let $\phi = (1+\sqrt{5})/2 \approx 1.618$ and $\widehat{\phi} = (1-\sqrt{5})/2 \approx -0.618$. If $F(n)$ satisfies the Fibonacci recurrence $F(n) = F(n-1) + F(n-2)$, we said in the text that $F(n) = \Theta(\phi^n)$. Let us now give the exact solution for this recurrence.
(a) Use induction to show that $F(n) = \phi^n/\sqrt{5} - \widehat{\phi}^n/\sqrt{5}$ is the solution with the initial

[2] Outside of the complexity of algorithms, we sometimes solve recurrences more accurately than its $\Theta$-order. E.g., $\mu(h) = \mu(h-1) + \mu(h-2) + 1$ for minimum size AVL trees (Chapter III). Even for complexity of algorithms, sharp bounds may be warranted: in the comparison model, we derive sharp bounds for sorting $S(n)$ or median $M(n)$ (Chapter I).

[3] Note that (5) is a compact way of saying that

$$F(0) = 0, F(1) = 1, F(2) = 1, F(3) = 2, F(4) = 3, \ldots.$$

We will use this convention throughout the book.

conditions $F(n) = n$ for $n = 0, 1$.
(b) Some authors like to begin with $F(n) = 1$ for $n = 0, 1$. Find the constants $a, b$ such that $F(n) = a\phi^n + b\widehat{\phi}^n$ for all $n \in \mathbb{N}$.
(c) Is it true that for all $n$ large enough, the standard Fibonacci sequence satisfies $F(n) = \lfloor a\phi^n \rceil$ for some constant $a$? Is it true if $F(n)$ is a "general" Fibonacci sequence in the sense that the initial values $F(0)$ and $F(1)$ are arbitrary (be careful). $\diamondsuit$

**Exercise 1.5:** (Fast Fibonacci Algorithm) For $n \geqslant 1$, let $P_n = \begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix}$ be the $n$-th pair of Fibonacci numbers, viewed as a column vector. Let $N = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ be the "next matrix".

(a) Show that $P_{n+1} = N \cdot P_n$ (matrix-vector multiplication).

(b) Conclude that $P_{n+1} = N^n \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

(c) Please compute $N^n$ for $n = 1, 2, 4, 8$.

(d) Compute $N^{11}$ using the matrices $N^8, N^4, N^2, N$ computed in the previous part.

(e) Prove that $N^n = \begin{bmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{bmatrix}$ for $n \geqslant 1$

(f) Make the case that this development will go more naturally if we extend our definition of $F_n$ to $F_{-1} = 1$ and use the recursive formula $F_n = F_{n-1} + F_{n-2}$ for all $n \geqslant 1$.

(g) Give a pseudo-code for computing $F_n$ using $O(\log n)$ arithmetic operations (addition or multiplication of integers).

$\diamondsuit$

**Exercise 1.6:** In the previous exercise (Fast Fibonacci Algorithm), we encounter the "addition chain" problem: given an integer $n > 1$, what is shortest **addition chain**

$$(a_1, \ldots, a_\ell) = (1, \ldots, n)$$

that starts with $a_1 = 1$ and ends with $a_\ell = n$. "Shortest" means $\ell$ is minimal. Let $\ell(n)$ denote this minimal value. In general, each $a_i$ $(i > 1)$ is obtained by adding two previous values. Thus $a_i = a_j + a_k$ where $1 \leqslant j \leqslant k < i$. Note that we allow $j = k$.

(a) Say how this addition chain problem can be used in the Fast Fibonacci Algorithm.

(b) Determine $\ell(n)$ for $n = 2, 3, 4, 5$.

(c) Suppose $\nu(n)$ is the number of 1's in the binary notation for $n$ (e.g., $\nu(2) = 1$ and $\nu(5) = 2$). Show that

$$\lg n + \lg \nu(n) - 3 \leqslant \ell(n) \leqslant \lg n + \nu(n)$$

REMARK: this problem can be generalized to the "addition sequence problem", where we want the shortest addition chain that contains a given set $\{n_1, \ldots, n_m\}$ of integers. See **Computing Sequences with Addition Chains**, P.Downey, B.Leong and R.Sethi. SIAM J.Comp. Vol.10, No.3, 1981, pp.638–646. This problem is NP-complete. $\diamondsuit$

**Exercise 1.7:** Let $T(n) = aT(n/b) + n$, where $a > 0$ and $b > 1$. How sensitive is this recurrence to the initial conditions? More precisely, if $T_1(n)$ and $T_2(n)$ are two solutions corresponding to two initial conditions, what is the strongest relation you can infer between $T_1$ and $T_2$? $\diamondsuit$

**Exercise 1.8:** (Aho and Sloane, 1973) Consider recurrences of the form

$$T(n) = (T(n-1))^2 + g(n). \tag{6}$$

For this exercise, assume $n$ is a natural numbers and use explicit boundary conditions (not DIC).
(a) Show that the number of binary trees of height at most $n$ is given by this recurrence with $g(n) = 1$ and the boundary condition $T(1) = 1$. Show that this particular case of (6) has solution

$$T(n) = \left\lfloor k^{2^n} \right\rfloor. \tag{7}$$

(b) Show that the number of Boolean functions on $n$ variables is given by (6) with $g(n) = 0$ and $T(1) = 2$. Solve this. ♢

**Exercise 1.9:** Let $T, T'$ be binary trees and $|T|$ denote the number of nodes in $T$. Define the relation $T \sim T'$ recursively as follows: (BASIS) If $|T| = 0$ or $1$ then $|T| = |T'|$. (INDUCTION) If $|T| > 1$ then $|T'| > 1$ and either (i) $T_L \sim T'_L$ and $T_R \sim T'_R$, or (ii) $T_L \sim T'_R$ and $T_R \sim T'_L$. Here $T_L$ and $T_R$ denote the left and right subtrees of $T$.
(a) Use this to give a recursive algorithm for checking if $T \sim T'$.
(b) Give the recurrence satisfied by the running time $t(n)$ of your algorithm.
(c) Give asymptotic bounds on $t(n)$. ♢

End Exercises

## §2. Simplification

**¶4.** We began with the Fibonacci recurrence (1) and Mergesort recurrence (2). They are representatives of two distinct areas of application: Mergesort-type recurrences arise in the analysis of recursive algorithms, and typically, they are polynomial in $n$. Fibonacci-type recurrences arise in combinatorial analysis (e.g., they arise in Chapter III when we analyze heights of certain search trees), and their solution tend to grow exponentially in $n$. In such analysis, we often get an infinite sequences of integers,

$$a_0, a_1, a_2, \ldots \tag{8}$$

where each $a_n$ represents the number of combinatorial objects of size $n$. For instance, $a_n$ may be the number of graphs on the vertex set $\{1, 2, \ldots, n\}$. Such sequences typically satisfy some recurrence equation. One of the simplest type of recurrence relations is the **linear recurrence** of the following form:

$$a_n = c + b_1 a_{n-1} + b_2 a_{n-2} + \cdots + b_d a_{n-d} = c + \sum_{i=1}^{d} b_i a_{n-i} \tag{9}$$

for all $n \geqslant d$. Here $d, c, b_i$ are constants. We call $d$ the **order** of this linear recurrence. If $c = 0$, we say the linear recurrence is **homogeneous** and otherwise **non-homogeneous**. Thus, the Fibonacci sequence is homogeneous of order 2. The theory of linear recurrences is well-understood; so we are contented with some Exercises and refer students to standard texts in combinatorial analysis. This chapter will largely focus on generalizations of Mergesort-type recurrences.

In the real world, actual recurrences can be rather messy. But since we are interested in Θ-order solutions, there are usually some simplifications steps to be taken.

**¶5. General Simplifications.** Here are three simplifications that should be automatically taken.

*taking a cue from Einstein in the epigraph...*

- **Initial Condition.** In this book, we normally state recurrences *without* initial conditions. In this case, we expect the student to supply the initial conditions of the following form:

  *DIC for convenience*

  > **Default Initial Condition** (DIC):
  > *There is some* $n_1 \geqslant 0$ *and* $C \geqslant 0$ *such that*
  > (1) *the recurrence for* $T(n)$ *holds for* $n > n_1$,
  > (2) $0 \leqslant T(n) \leqslant C$ *for* $0 \leqslant n \leqslant n_1$. (10)

  Of course, the justification for DIC is that, up to Θ-order, it has no effect on the solution. Note that DIC is a scheme for a large class of initial conditions. Even if you fix $n_1$ and $C$, there are still infinitely many initial values for $T(n)$ based on $n_1, C$. Our favorite form of DIC is the **constant DIC**, namely, there is some constant $C \geqslant 0$ such that $T(n) = C$ for all $0 \leqslant n \leqslant n_1$. This DIC is uniquely determined by $n_1$ and $C$. Since this is so simple, why would you use other forms of DIC? One answer is that other DICs may lead to simpler forms of the solution. For instance, we could let $T(n) = C(n)$ for all $n < n_0$ where $C(n)$ some a function of $n$. See ¶8 below for an example. In using DIC, we need not specify $n_1$ or the initial values of $T(n)$ in advance: instead, we can just proceed to solve the recurrence and, at the appropriate moments, introduce these values.

  *Thus DIC allows us to focus on the recurrence itself rather than the initial conditions.* We had assumed that DIC does not affect the asymptotic solution; but even if DIC could the solution, we may have learned something about the recurrence. We have seen in the Fibonacci recurrence, a particular initial condition could lead to the trivial solution $F(n) \equiv 0$. What other behaviors can arise for $F(n)$? We could get $F(n) < 0$ for all $n$ large enough (see Exercise). But our DIC excludes this possibility. This is enough to ensure that all solutions have the same Θ-order.

- **Extension to Real Functions.** Even if the function $T(n)$ is originally defined for natural numbers $n$, we will now treat $T(n)$ as a real function (*i.e.*, $n$ is viewed as a real variable), and defined for $n$ sufficiently large. In the case of integer sequences, it is unclear what the real analogue of (8) really means; nevertheless, their recurrence relation (9) is meaningful when the index $n$ is treated as a real number. The traditional approach is to solve the recurrence on an "ample domain" (see Exercises) in order to avoid extensions into real functions. For example, in the Mergesort Recurrence (2), to solve for $T(n)$, we may be forced to define $T(x)$ for non-integer $x$ after some recursive expansion. But we note that we can avoid non-integer $x$, if $n$ is a power of 2. We call the set $D = \left\{2^k : k \in \mathbb{N}\right\}$ an "ample domain" for this recurrence. After solving $T(n)$ for all $n \in D$, we can now bound $T(n)$ when $n \notin D$ by some smoothness or monotonicity assumption on $T(n)$. It is important to realize that even if we have no interest in real recurrences, some solution techniques below will transform our recurrences into non-integer recurrences. So we might as well take the plunge from the start. But the best recommendation for our real approach is its naturalness and the resulting simplicity.

- **Converting Recurrence Inequality into a Recurrence Equation.** If we begin with a recurrence inequality such as $T(n) \leqslant G(n, T(n_1), \ldots, T(n_k))$, we simply rewrite this as an equality relation: $T(n) = G(T(n_1), \ldots, T(n_k))$. Because of this change, our eventual solution for $T(n)$ is only an upper bound on the original function. Similarly, if we had started with $T(n) \geqslant G(n, T(n_1), \ldots, T(n_k))$, the eventual solution is only a lower bound.

**¶6. Special Simplifications.** Suppose the running time of an algorithm satisfies the following inequality:

$$T(n)\left\{\begin{array}{cll} \leqslant & T(\lceil n/2\rceil)+T(\lfloor n/2\rfloor)+6n+\lg n-4 & \text{for integer } n>100,\\ = & 3n^2-4n+2 & \text{for } 0\leqslant n\leqslant 100. \end{array}\right. \tag{11}$$

Such a **recurrence inequality** might arise in some imagined implementation of Mergesort, with special treatment for $n \leqslant 100$. Our general simplification procedure tells us to (a) discard the specific boundary conditions (for $0 \leqslant n \leqslant 100$) in favor of DIC, and (b) treat $T(n)$ as a real function, and (c) write the recurrence as a equation. This leaves us with

$$T(n) = T(\lceil n/2\rceil) + T(\lfloor n/2\rfloor) + 6n + \lg n - 4. \tag{12}$$

What further simplifications might apply here? We can further convert (12) into

$$T(n) = 2T(n/2) + n. \tag{13}$$

This represents two additional simplifications: (i) We replaced the term "$+6n + \lg n - 4$" by some simple expression ("$+n$") with the same $\Theta$-order. (ii) We have removed the ceiling and floor functions. Step (i) is justified because this does not affect the $\Theta$-order (if this is not clear, then you can always come back to verify this claim). Step (ii) exploits the fact that we now treat $T(n)$ as a real function, so we need not worry about non-integral arguments when we remove the ceiling or floor functions. Also, it does not affect the asymptotic value of $T(n)$ here.

The justifications for these steps are certainly not obvious, but they should seem reasonable. Ultimately, one ought to return to such simplifications to justify them.

### Exercises

**Exercise 2.1:** Show that our above simplifications of the the recurrence (11) (with its initial conditions) cannot affect the asymptotic order of the solution. [Show this for ANY choice of Default Initial Condition.] ♢

**Exercise 2.2:** We seek counter-examples to the claim that we can replace $\lceil n/2\rceil$ by $n/2$ in a recurrence without changing the $\Theta$-order of the solution.

(a) Construct a function $g(n)$ that provides a counter example for the following real recurrence: $T(n) = T(\lceil n/2\rceil) + g(n)$. HINT: make $g(n) = 1$ iff $n$ is an odd integer, 0 otherwise.

(b) Construct a different counter example of the form $T(n) = h(n)T(\lceil \frac{n}{2}\rceil)$ for a suitable function $h(n)$. HINT: make $h(n)$ grow very fast.

♢

**Exercise 2.3:** Show examples where the choice of initial conditions can change the $\Theta$-order of the solution $T(n)$. NOTE: We already know that this can happen when the initial condition leads to the trivial solution, so avoid this case. HINT: Choose $T(n)$ to increase exponentially. ♢

**Exercise 2.4:** Suppose $x, n$ are positive numbers satisfying the following "non-separable recurrence" equation,

$$2^x = x^{2n}.$$

Solve for $x$ as a function of $n$, showing

$$x(n) = [1 + o(1)]2n\log_2(2n).$$

HINT: take logarithms. This is an example of a bootstrapping argument where we use an approximation of $x(n)$ to derive yet a better approximation. See, e.g., Purdom and Brown [16]. ♢

**Exercise 2.5:** [Ample Domains] Our approach of considering real functions is non-standard. The more typical approach in the algorithms literature is as follows. Consider the simplification of (11) to (13). Suppose, instead of assuming $T(n)$ to be a real function (so that (13) makes sense for all values of $n$), we continue to assume $n$ is a natural number. It is easy to see that $T(n)$ is completely defined by (13) iff $n$ is a power of 2. We say that (13) is closed over the set $D_0 := \{2^k : k \in \mathbb{N}\}$ of powers of 2. In general, we say a recurrence is "closed over a set $D \subseteq \mathbb{R}$" if for all $n \in D$, the recurrence for $T(n)$ depends only on smaller values $n_i$ that also belong in $D$ (unless $n_i$ lies within the boundary condition).
(a) Let us call a set $D \subseteq \mathbb{R}$ an "ample set" if, for some $\alpha > 1$, the set $D \cap [n, \alpha \cdot n]$ is non-empty for all $n \in \mathbb{N}$. Here $[n, \alpha n]$ is closed real interval between $n$ and $\alpha n$. If the solution $T(n)$ is sufficiently "smooth", then knowing the values of $T(n)$ at an ample set $D$ gives us a good approximation to values where $n \notin D$. In this question, our "smoothness assumption" is simply: *$T(n)$ is monotonic non-decreasing.* Suppose that $T(n) = n^k$ for $n$ ranging over an ample set $D$. What can you say about $T(n)$ for $n \notin D$? What if $T(n) = c^n$ over $D$? What if $T(n) = 2^{2^n}$ over $D$?
(b) Suppose $T(n)$ is recursively expressed in terms of $T(n_1)$ where $n_1 < n$ is the largest prime smaller than $n$. Is this recurrence defined over an ample set? ♢

**Exercise 2.6:** Consider inversions in a sequence of numbers.
(a) The sequence $S_0 = (1, 2, 3, 4)$ has no inversions, but sequence $S_1 = (2, 1, 4, 3)$ has two inversions, namely the pairs $\{1, 2\}$ and $\{3, 4\}$. Now, the sequence $S_2 = (2, 3, 1, 4)$ also has two inversions, namely the pairs $\{1, 2\}$ and $\{1, 3\}$. Let $I(S)$ be the number of inversions in $S$. Give an $O(n \lg n)$ algorithm to compute $I(S)$. Hint: use a generalization of Mergesort.
(b) We next distinguish between the quality of the inversions of $S_1$ and $S_2$. The inversions $\{1, 2\}$ and $\{3, 4\}$ in $S_1$ are said to have weight of 1 each, so the **weighted inversion** of $S_1$ is $W(S_1) = 2 = 1 + 1$. But for $S_2$, the inversion $\{1, 2\}$ has weight 2 while inversion $\{1, 3\}$ has weight 1. So the weighted inversion is $W(S_2) = 3 = 2 + 1$. Thus the "weight" measures how far apart the two numbers are. In general, if $S = (a_1, \ldots, a_n)$ then a pair $\{a_i, a_j\}$ is an **inversion** if $i < j$ and $a_i > a_j$. The weight of this inversion is $j - i$. Let $W(S)$ be the sum of the weights of all inversions. Give an $O(n \lg n)$ algorithm for weighted inversions. ♢

**Exercise 2.7:** We might consider following form of DIC where we assume that there exists $0 < n_0 < n_1$, and constants $0 < C_0 \leqslant C_1$ such that

$$(\forall\ n_0 \leqslant n < n_1)[C_0 \leqslant T(n) \leqslant C_1]. \tag{14}$$

Solve the Fibonacci and mergesort recurrences using this version of DIC. Your solutions should be stated in terms of the parameters $C_1, C_2$. ♢

**Exercise 2.8:** Consider the Fibonacci recurrence $F(n) = F(n-1) + F(n-2)$. Let $DIC(C, n_0)$ denote the initial condition that $F(n) = C$ for $n \leqslant n_0$. Assume $C > 0$ and $n_0 > 1$. For any $k$, let $F_k(n)$ be the solution to the Fibonacci recurrence under the initial condition $DIC(k, k)$. Prove that $F_k(n) = \Theta(F_j(n))$ for any $k, j > 1$. ♢

**Exercise 2.9:** Associated to the linear recurrence (9) is the **characteristic polynomial**

$$x^d = b_1 x^{d-1} + b_2 x^{d-2} + \cdots + b_d = \sum_{i=1}^{d} b_i x^{d-i} \tag{15}$$

The $d$ roots $\lambda_1, \ldots, \lambda_d$ of this polynomial are called the **characteristic roots**. They are not necessarily distinct and can be complex. In this question, assume that the $\lambda_i$'s are all distinct; also assume that $c = 0$ (recurrence is homogeneous).

(a) Show that for all initial values $a_0, \ldots, a_{d-1}$ of the recurrence sequence, there exists constants $c_1, \ldots, c_d$ such that the recurrence sequence satisfies

$$a_n = \sum_{i=1}^{d} c_i \lambda_i^n \tag{16}$$

for all $n \geqslant 0$.

HINT: best to use matrix notations. Let $V_n = \begin{bmatrix} \lambda_1^{n-d} & \cdots & \lambda_1^{n-1} \\ \vdots & \ddots & \vdots \\ \lambda_d^{n-d} & \cdots & \lambda_d^{n-1} \end{bmatrix}$ for all $n \geqslant d$. For instance, $V_d$ is usually called the Vandermonde matrix. It is well-known that $V_d$ is non-singular when the $\lambda_i$'s are distinct. Then, by definition of $\lambda_i$, we have

$$\begin{bmatrix} \lambda_1^d \\ \vdots \\ \lambda_d^d \end{bmatrix} = V_d \begin{bmatrix} b_d \\ \vdots \\ b_1 \end{bmatrix}. \tag{17}$$

(b) Suppose there are only $m$ distinct roots where $m < d$, say $\lambda_1, \ldots, \lambda_m$. Let each $\lambda_i$ has multiplicity $\mu_i \geqslant 1$ with $\mu_1 + \cdots + \mu_m = d$. Show that the general solution is given by

$$a_n = \sum_{i=1}^{m} \sum_{j=1}^{\mu_i} c_{i,j} \lambda_i^{n-j} \tag{18}$$

where the $d$ coefficients $c_{i,j}$'s are again determined by the initial conditions.

(c) Let us apply your solution for part (a) to the Fibonacci recurrence $F(n)$, except that the initial values are $F(0) = u$ and $F(1) = v$ for some variables $u, v$. Determine the coefficients $c_1, c_2$ in (16). Thus each $c_i = c_i(u, v)$ is a function of $u, v$.

(d) Let $a_n = a_{n-1} + a_{n-2} + a_{n-3}$. Determine its characteristic roots. If $a_i = i$ for $i = 0, 1, 2$, give the general formula for $a_n$ $(n \geqslant 3)$.

(e) Find the general formula for $a_n$ when the linear recurrence has characteristic polynomial $p(x) = x^3 - 2x^2 - 5x + 6$ and the initial conditions are $a_n = n$ for $n = 0, 1, 2$.

◇

End Exercises

## §3. Divide-and-Conquer Algorithms

In this section, we look at an interesting type of recurrence that arises in many divide-and-conquer algorithms. We call it the "Master Recurrence" because the well-known **Master Theorem** is the solution for this this recurrences. We also look at two concrete algorithms of this kind: Karatsuba's classic algorithm (1962) for multiplying integers [10], and a modern problem arising in searching for key words.

**¶7. Master Recurrence and Divide-and-Conquer Algorithms.** The recurrences (2) is an instance of the **Master Recurrence** which has the form:

$$T(n) = aT(n/b) + d(n) \tag{19}$$

where $a > 0$ and $b > 1$ are real constants and $d$ is any function, called[4] the **driving** or **forcing function**. We shall solve this recurrence under fairly general conditions: for instance, the literature sometimes assume $a$ and $b$ are integers but we have no such restrictions. The following table show some $a, b, d(n)$ arising in actual algorithms:

| **Algorithm** | $a$ | $b$ | $d(n)$ |
|---|---|---|---|
| Mergesort | 2 | 2 | $n$ |
| Karatsuba integer multiplication | 3 | 2 | $n$ |
| Strassen matrix multiplication | 7 | 2 | $n^2$ |
| Pan matrix multiplication (1978) | 143640 | 70 | $n^2$ |

The idea of solving a problem by reducing it to smaller subproblems is very general. We mainly focus on reductions from problems of size $n$ to subproblems of size $\leqslant n/b$ for one or more constants $b > 1$. In other problems, we reduce a problem of size $n$ to several subproblems that of size $\leqslant n - c$ for some fixed $c \geqslant 1$. Such solutions would be exponential time without additional properties; we study these under the topic of dynamic programming (Chapter 7). In applications, we have $d(n) > 0$, representing the cost of merging solutions of subproblems in divide-and-conquer algorithms.

*OK, some of you learned it in grade school*

**¶8. Example from Arithmetic.** To motivate Karatsuba's algorithm, let us recall the classic "high-school algorithm" for multiplying integers. Given positive integers $X, Y$, we want to compute their product $Z = XY$. This algorithm assumes you know how to do single-digit multiplication and multi-digit additions ("pre-high school"). The algorithm multiplies $X$ by each digit of $Y$. If $X$ and $Y$ have $n$ digits each, then we now have $n$ products. E.g., with $n = 3$, $X = 123$ and $Y = 789$ then we have the products $7 \cdot X = 861$, $8 \cdot X = 984$, $9 \cdot X = 1{,}087$. Each of these products has at most $n + 1$ digits. After appropriate left-shifts of these $n$ products, we add them up to get the final product: $123 \times 789 = 700X + 80X + 9X = 86{,}100 + 9{,}840 + 1{,}087 = 97{,}047$. It is not hard to see that this algorithm takes $\Theta(n^2)$ time. Can we improve on this?

*No Roman numerals, please. See the epigraph of Dijkstra in this chapter.*

Usually we think of $X, Y$ in decimal notation, but the algorithm works equally well in any base. We shall assume base 10 in our hand simulations. For instance, if $X = 19$ (nineteen) then in binary $X = 10011$. To avoid the ambiguity from different bases, we indicate[5] the base using a subscript, $X = (10011)_2$. By convention, the decimal base is assumed when no base is indicated. Thus a plain "100" without any base represents one hundred, but $(100)_2$ represents four.

Assume $X$ and $Y$ has length exactly $n$ where $n$ is a power of 2. This is without loss of generality since we can always pad with $X$ and $Y$ with leading 0's. In our above example, $X = 123$ is written 0123 and $Y = 0789$ of length $n = 4 = 2^k$ ($k = 2$). Next, we split up $X$ into a high-order half $X_1$ and low-order half $X_0$. If $X = 0123$ then $X_1 = 01$ and $X_0 = 23$. Thus

$$X = X_0 + 10^{n/2} X_1$$

[4] The presence of a non-zero driving function $d(n)$ makes this recurrence "non-homogeneous". But for recurrences from complexity of algorithms, we even have $d(n) > 0$.

[5] By the same token, we may write $X = (19)_{10}$ for base 10. But now the base "10" itself may be ambiguous — after all "10" in binary is equal to two. By convention we write the base in decimal.

where $X_0, X_1$ are $n/2$-bit numbers. Similarly, $Y = Y_0 + 10^{n/2}Y_1$ and

$$\begin{aligned} Z &= (X_0 + 10^{n/2}X_1)(Y_0 + 10^{n/2}Y_1) \\ &= X_0Y_0 + 10^{n/2}(X_1Y_0 + X_0Y_1) + 10^n X_1Y_1 \\ &= Z_0 + 10^{n/2}Z_1 + 10^n Z_2, \end{aligned}$$

where $Z_0 = X_0Y_0$, etc. Clearly, each of these $Z_i$'s have at most $2n$ bits. Now, if we compute the 4 products

$$X_0Y_0, X_1Y_0, X_0Y_1, X_1Y_1$$

recursively, then we can put them together ("conquer step") in $\mathcal{O}(n)$ time. To see this, we must make an observation: in binary notation, multiplying any number $X$ by $2^k$ (for any positive integer $k$) takes $\mathcal{O}(k)$ time, independent of $X$. We can view this as a matter of shifting left by $k$, or by prepending a string of $k$ zeros to $X$.

Hence, if $T(n)$ is the time to multiply two $n$-bit numbers, we obtain the recurrence

$$T(n) \leqslant 4T(n/2) + Cn \tag{20}$$

for some $C > 1$. Given our guidelines for simplification of recurrences, we immediately rewrite this as

$$T(n) = 4T(n/2) + n.$$

As we will see, this recurrence has solution $T(n) = \Theta(n^2)$. So we have not really improved upon the high-school method!

Karatsuba (1962) observed that we can proceed as follows: first compute $Z_0 = X_0Y_0$ and $Z_2 = X_1Y_1$ first. Then we can compute $Z_1$ with only one multiplication and four additions/subtractions:

$$Z_1 = (X_0 + X_1)(Y_0 + Y_1) - Z_0 - Z_2.$$

Thus $Z_1$ can be computed with one recursive multiplication plus some additional $\mathcal{O}(n)$ work. From $Z_0, Z_1, Z_2$, we can again obtain $Z$ in $\mathcal{O}(n)$ time. This gives us the **Karatsuba recurrence**,

$$T(n) = 3T(n/2) + n. \tag{21}$$

We shall show that $T(n) = \Theta(n^\alpha)$ where $\alpha = \lg 3 = 1.58\cdots$. This is clearly an improvement of the high school method.

*First improvement in 1000 years? Wikipedia says the high school multiplication is equivalent to the "lattice method" which is at least 1000 years old.*

Reality check: if $X, Y$ have different bit lengths, we can "pad" one of them so that they have the same length $n$. But in recursive calls, one of the half may have length $\lceil n/2 \rceil$. Moreover, the product $Z_1 = (X_0+X_1)(Y_0+Y_1)$ involve may involve $\lceil n/2 \rceil + 1$ bit numbers. So the Karatsuba's recurrence is more accurately given by

$$T(n) \leqslant 3T(\lceil n/2 \rceil + 1) + Cn$$

In an exercise, we will show that none of these perturbations affect our Theta-bound.

**Hand-Simulation:** We want students to do hand-simulation to really understand Karatsuba's algorithm. The method we suggest is to organize a "up-down recursion tree": the usual recursion tree tells us how to split the problem into subproblems: at each node $u$, we store the 2 integers $(u_X, u_Y)$ which are the arguments for multiplication. This in tern generates three children $u_0, u_1, u_2$ with their recursive arguments. This represents "going down" the recursion tree. In "going up" the recursion tree, want to store the product $u_Xu_Y$ at node $u$. This is obtained from the products at the children $u_0, u_1, u_3$. One should do some side computation to reconstruct $u_Xu_Y$. E.g., Figure 1 shows such a simulation:

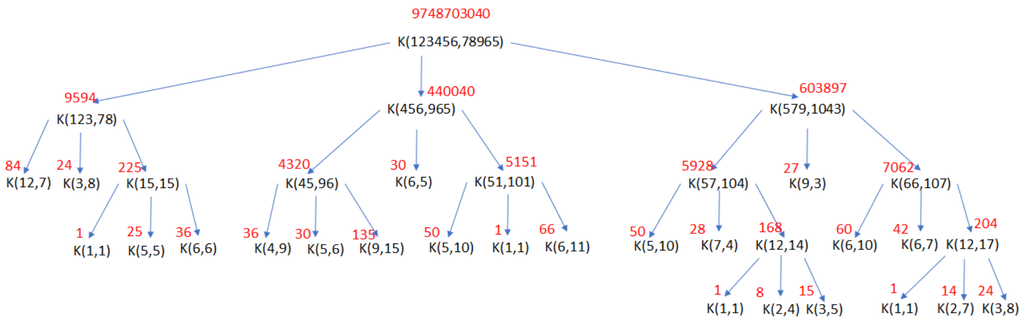


Figure 1: Karatsuba simulation for 123456 × 78965

**NOTE:** There is an even faster multiplication algorithm from Schönhage and Strassen (1971) that runs in time $\mathcal{O}(n \log n \log\log n)$. There is an increasing need for multiplication of arbitrarily large integers. In cryptography or computational number theory, for example. These are typically implemented in software in a "big integer" package. For instance, `Java` has a `BigInteger` class. A well-engineered big integer multiplication algorithm will typically implement the High-School algorithm for $n \leqslant n_0$, and use Karatsuba for $n_0 < n \leqslant n_1$, and use Schönhage-Strassen for $n > n_1$. Typical values for $n_0, n_1$ are $30, 200$ digits. One of the oldest questions in theoretical computer science concerns the inherent complexity of multiplication. In particular, is $O(n \log n \log\log n)$ the best possible? Most computer scientists believe that $O(n \log n)$ is the right answer. After more than 30 years, finally M. Fürer (2007) breached the $\log\log n$ factor. He achieved an $O(n \log n \log^* n)$ multiplication algorithm. In 2008, A. De, C. Saha, P. Kurur and R. Saptharish achieved the same bound by a different method, based on modular arithmetic.

**¶9. A Google Problem.** The Google Phenomenon is possible because of efficient algorithms: every file on the web can be searched and indexed. Searching is by keywords. E.g., Find me all files with the keyword 'algorithm'. Clearly, we need to filter this by some contextual information to get a meaningful output. Let us suppose that Google pre-processes every file in its database for keywords. Of course there are many issues, even in this innocent task: should we distinguish upper/lower cases ('Algorithm' versus 'algorithm'), variant spelling, grammatical variants (plural/singular, etc), and even wrongly spelled words ('algorthm'). But let just assume that there is a single keyword that is abstractly denoted [algorithm], which may or may not include variants or even multilingual forms. We are interested in searching based on more than one keyword: E.g. Find me all files with the keywords [algorithm] and [introduction]. We will reduce this multi-keyword search to a precomputed single-keyword index.

Let $F$ be a file, viewed as a sequence of words (ignoring punctuation, capitalization, etc). We first pre-process $F$ for the occurrences of keywords. For each keyword $w$, we precompute an **index** which amounts a sorted sequence $P(w)$ of positions indicating where $w$ occurs in $F$. E.g.,

$$P(divide) = (11, 16, 42, 101, 125, 767)$$

means that the keyword *divide* occurs in $F$ at positions $11, 16$, etc, for a total of 6 times. Suppose we want to search the file using a conjunction of $k$ keywords, $w_1, \ldots, w_k$. An interval $J = [s, t]$ is called a **cover** for $W = \{w_1, \ldots, w_k\}$ if each $w_i$ occurs at least once within the positions in $J$. The size of a cover $[s, t]$ is just $t - s$. A cover is **minimal** if it does not contain any smaller cover; it is **minimum** if its size is smallest among all covers. The **keyword cover problem** is this: given a set $W = \{w_1, \ldots, w_k\}$ of key words, and also indices $P(w_1), \ldots, P(w_k)$ for these key words in a file, to compute a minimum cover for $W$.

E.g., let $k = 2$ with $w_1 = divide$ and $w_2 = conquer$. With $P(divide)$ as before, suppose $P(conquer) = (2, 44, 289, 300)$. Then the minimal covers are $[2, 11], [42, 44], [44, 101], [125, 289], [300, 767]$. This is illustrated in Figure 2. The minimum cover is $[42, 44]$.

*you think the word 'and' is at position 43?*

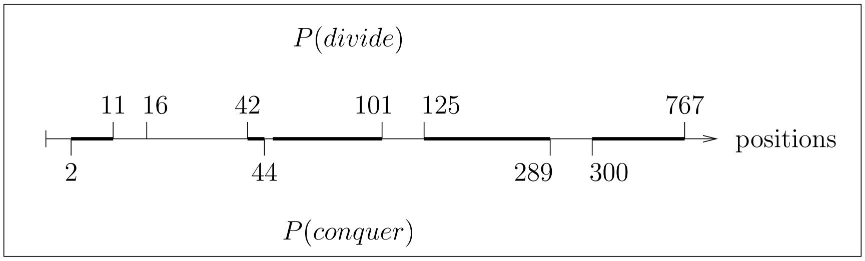


Figure 2: Minimal Covers

Before attempting to solve this problem, consider how Google might use the minimum cover solutions: suppose a user wants to search for a set $W = \{w_1, \ldots, w_k\}$ of key words. For each file $f_j$ $(j = 1, 2, \ldots)$ we use the algorithm to compute a minimum cover $[c_j, d_j]$ (if one exists) for $W$ in $f_j$. The indices $P(w_i)$ for each key word $w_i$ are assumed to have been precomputed. The search results will be a list of all files for which covers exist, but we order these files in order of non-decreasing cover size $d_j - c_j$. The actual cover $[c_j, d_j]$ can be used by Google to display a snippet of the file $f_j$.

Let us now consider algorithms. Let $n_i$ be the length of list $P(w_i)$ $(i = 1, \ldots, k)$ and $n = n_1 + \cdots + n_k$. The case $k = 2$ is relatively straightforward, and we leave it for an exercise. Consider the case $k = 3$. First, merge $P(w_1), P(w_2), P(w_3)$ into the array $A[1..n]$. Recall that in Chapter I, we discussed the merging of sorted lists. Merging takes time $O(n_1 + n_2 + n_3) = O(n)$. To keep track of the origin of each number in $A$, we may also construct an array $B[1..n]$ such that $B[i] = j \in \{1, 2, 3\}$ iff $A[i]$ comes from the list $P(w_j)$.

We use a divide-and-conquer approach. Recursively, compute a minimum cover of $A[1..(n/2)]$ and $A[(n/2) + 1..n]$ (for simplicity, assume $n$ is a power of 2). Let $C_{1,n/2}$ and $C_{(n/2)+1,n}$ be these minimum covers. We now need to find a minimal cover that straddles $A[(n/2)]$ and $A[(n/2) + 1]$. Let $C = [A[i], A[j]]$ be such a minimal cover, where $i \leqslant (n/2)$ and $j \geqslant (n/2) + 1$. There are 6 cases. One case is when $C = C' \cup C''$, where $C' = [A[i], A[n/2]]$ is the rightmost cover for $w_1$ in $A[1..(n/2)]$, and $C'' = [A[(n/2) + 1], A[j]]$ is the leftmost cover for $w_2, w_3$ in $A[(n/2) + 1, n]$. We can find $C'$ and $C''$ in $O(n)$ time. The remaining 5 cases can similarly be found in $O(n)$ time. Then $C$ is the cover that has minimum size among these 6 cases. Hence, the overall complexity of the algorithm satisfies

$$T(n) = 2T(n/2) + n.$$

We have seen this recurrence before, as the Mergesort recurrence (2). The solution is $T(n) = \Theta(n \log n)$. See exercise for a general solution in $O(n \log k)$ time.

## Exercises

**Exercise 3.1:** Carry out Karatsuba's algorithm for $X = 6 = (0110)_2$ and $Y = 11 = (1011)_2$. Draw the "up-down" recursion tree as described in the text. Notes: Assume $|X| = |Y| = n$, and when we split the argument, assume $|X_0| = \lceil n/2 \rceil$ and $|X_1| = \lfloor n/2 \rfloor$. Each internal node has 3 children corresponding to the results $Z_2, Z_1, Z_0$ (in this order please). ♢

**Exercise 3.2:** Suppose an implementation of Karatsuba's algorithm achieves $T(n) \leqslant Cn^{1.58}$ where $C = 1000$. Moreover, the High School multiplication is $T(n) = 30n^2$. Beyond what value of $n$ does Karatsuba definitely becomes competitive with the High School method? ♢

**Exercise 3.3:** Consider the recurrence $T(n) = 3T(n/2) + n$ and $T'(n) = 3T'(\lceil n/2 \rceil) + Cn$ (for some constant $C > 1$). Show that $T(n) = \Theta(T'(n))$. Thus, the presence of $\lceil \cdot \rceil$ and $C$ in $T'$ did not make it asymptotically different from $T$. HINT: Use the fact that $\lceil \lceil n/2^i \rceil /2 \rceil = \lceil n/2^{i+1} \rceil$. ♢

| NumBits | AvgTime | Exponent |
|---|---|---|
| 4000 | 4.358 | 0.0 |
| 4200 | 4.696 | 1.531002145103799 |
| 4400 | 5.194 | 1.841260577604784 |
| 4600 | 5.517 | 1.6873048110254347 |
| 4800 | 5.983 | 1.7381865504999572 |
| 5000 | 6.51 | 1.7985113947251763 |
| 5200 | 6.988 | 1.7997159663026001 |
| 5400 | 7.509 | 1.812998128928515 |
| 5600 | 8.01 | 1.8089977665618309 |
| 5800 | 8.684 | 1.85558837393382 |
| 6000 | 9.183 | 1.838236378924439 |
| 6200 | 9.769 | 1.8418523402197153 |
| 6400 | 10.365 | 1.8434357852847953 |
| 6600 | 11.088 | 1.864808884276074 |
| 6800 | 11.717 | 1.8638802969571109 |
| 7000 | 12.413 | 1.8704459319724756 |
| 7200 | 13.092 | 1.8714070696035303 |
| 7400 | 13.843 | 1.8787279477010768 |
| 7600 | 14.532 | 1.8763458534440565 |
| 7800 | 15.297 | 1.8801860861195574 |
| 8000 | 16.054 | 1.8811947011507577 |
| 8200 | 16.905 | 1.8884383570994894 |
| 8400 | 17.644 | 1.8847717474449632 |
| 8600 | 18.498 | 1.8885827751677746 |
| 8800 | 19.283 | 1.8862283707110576 |
| 9000 | 20.225 | 1.8927722703240168 |
| 9200 | 21.17 | 1.8976522229154338 |
| 9400 | 22.063 | 1.8982439890258536 |
| 9600 | 23.034 | 1.9017905239616146 |
| 9800 | 24.055 | 1.9064306092855452 |
| 10000 | 24.986 | 1.905838802838669 |
| 10200 | 25.987 | 1.9074840762036238 |
| 10400 | 26.948 | 1.9067232067781992 |
| 10600 | 28.108 | 1.912700793571853 |
| 10800 | 29.111 | 1.9120055203582398 |
| 11000 | 30.221 | 1.9143159996069712 |
| 11200 | 31.534 | 1.922120988851413 |
| 11400 | 31.542 | 1.8898795547030012 |
| 11600 | 32.67 | 1.8920105894497778 |
| 11800 | 33.703 | 1.8908891117429292 |
| 12000 | 34.67 | 1.8877101089855162 |
| 12200 | 36.082 | 1.8955269064390694 |
| 12400 | 37.218 | 1.8956825843907563 |
| 12600 | 38.049 | 1.8884930574030907 |
| 12800 | 39.242 | 1.8894663931349043 |
| 13000 | 40.553 | 1.892493164635265 |
| 13200 | 41.696 | 1.8915733844170872 |
| 13400 | 42.951 | 1.8925738155123988 |
| 13600 | 44.159 | 1.8923271871808227 |
| 13800 | 45.533 | 1.8947617307075215 |
| 14000 | 46.816 | 1.8951803717241376 |
| 14200 | 48.1 | 1.8953182704475686 |
| 14400 | 49.401 | 1.8954588786790316 |
| 14600 | 50.873 | 1.8979435636574864 |
| 14800 | 52.364 | 1.9002856600816482 |
| 15000 | 53.537 | 1.8977482007273088 |

Figure 3: Timing as a function of number of bits

**Exercise 3.4:** The following is a programming exercise. It is best done using a programming language such as Java that has a readily available library of big integers.
(a) Implement Karatsuba's algorithm using such a programming language and using its big integer data structures and related facilities. The only restriction is that you must not use the multiplication, squaring, division or reciprocal facility of the library. But you are free to use its addition/subtraction operations, and any ability to perform left/right shifts (multiplication by powers of 2).
(b) Let us measure the running time of your implementation of Karatsuba's algorithm. For input numbers, use a random number generator to produce numbers of any desired bit length. If $T(n) \leqslant Cn^{\alpha}$ then $\lg T(n) \leqslant \lg C + \alpha \lg n$. The **exponent** $\alpha$ is thus the slope of the curve obtained by plotting $\lg T(n)$ against $\lg n$, we should get a slope of at most $\alpha$. Plot the running time of your implementation to verify that its exponent is $< 1.58$.
(c) What is the exponent in Java's native implementation? Explain your data.
(d) My 1999 undergraduate class in algorithms did the preceding exercise, using the `java.math.BigInteger` package. One timing from this class is shown in Table 3. The "exponent" in this table is computing with a crude formula $\frac{\lg(avgTime)-avgTime_0}{\lg(numBits)-numBits_0}$ where $numBits_0 = 4000$ and $avgTime_0 = 4.358$ (the initial trial). This crude exponent hovers around 1.9. What would be the empirical exponent if you do a proper regression analysis? This data suggests that in 1999, the library only implemented the High School algorithm. By 2001, the situation appeared to have improved. ♢

**Exercise 3.5:** Suppose the running time of an algorithm is an unknown function of the form $T(n) = An^a + Bn^b$ where $a > b$ and $A, B$ are arbitrary positive constants. You want to discover the exponent $a$ by measurement. How can you, by plotting the running time of the algorithm for various $n$, find $a$ with an error of at most $\epsilon$? Assume that you can do least squares line fitting. ♢

**Exercise 3.6:** Try to generalize Karatsuba's algorithm by breaking up each $n$-bit number into 3 parts. What recurrence can you achieve in your approach? Does your recurrence improve upon Karatsuba's exponent of $\lg 3 = 1.58\cdots$? ♢

**Exercise 3.7:** To generalize Karatsuba's algorithm, consider splitting an $n$-bit integer $X$ into $m$ equal parts (assuming $m$ divides $n$). Let the parts be $X_0, X_1, \ldots, X_{m-1}$ where $X = \sum_{i=0}^{m-1} X_i 2^{in/m}$. Similarly, let $Y = \sum_{i=0}^{m-1} Y_i 2^{in/m}$. Let us define $Z_i = \sum_{j=0}^{i} X_j Y_{i-j}$ for $i = 0, 1, \ldots, 2m-2$. In the formula for $Z_i$, assume $X_\ell = Y_\ell = 0$ when $\ell \geqslant m$.
(i) Determine the Θ-order of $f(m,n)$, defined to be the time to compute the product $Z = XY$ when you are given $Z_0, Z_1, \ldots, Z_{2m-2}$. Remember that $f(m,n)$ is the number of bit operations.
(ii) It is known that we can compute $\{Z_0, Z_1, \ldots, Z_{2m-2}\}$ from the $X_i$'s and $Y_j$'s using $\mathcal{O}(m \log m)$ multiplications and $\mathcal{O}(m \log m)$ additions, all involving $(n/m)$-bit integers. Using this fact with part (i), give a recurrence relations for the time $T(n)$ to multiply two $n$-bit integers.
(iii) Conclude that for every $\varepsilon > 0$, there is an algorithm for multiplying any two $n$-bit integers in time $T(n) = \Theta(n^{1+\varepsilon})$. NOTE: part (iii) is best attempted after you have studied the Master Theorem in the subsequent sections. ◇

**Exercise 3.8:** In the Google problem, we need to merge several sorted lists. Recall from Chapter I that we can merge a two lists of sizes $n$ and $n'$ in time $\Theta(n + n')$. Suppose $X_1, \ldots, X_k$ are $k \geqslant 1$ sorted lists, each with $n \geqslant 1$ elements. Here, $k$ and $n$ are independent parameters.
(a) We want to analyze the complexity $T(n,k)$ of sorting the set $X = \bigcup_{i=1}^{k} X_i$. At each phase, we merge pairs of lists. With $k$ lists of size $n$, we take $O(nk)$ time to merge, and produce $k/2$ lists each of size $2n$. Set up the recurrence for $T(n,k)$ based on this repeated merging algorithm.
(b) Show that $T(n,k) = \mathcal{O}(nk \lg k)$) HINT: you could use domain transformation (see §7).
(c) Use the Information Theoretic Lower Bound from Chapter I to show a lower bound of $\Omega(nk \lg k)$. ◇

*Adapted from a Google interview question (the interviewed student Z. was hired)*

**Exercise 3.9:** Recall the Google multi-keyword search. This was reduced to computing a minimum cover for a set $W = \{w_1, \ldots, w_k\}$ of key words in a file. For each key word $w_i \in W$, we are given an index $P(w_i)$ which is just a sorted list of positions where $w_i$ occurs in the file. Let $n = \sum_{i=1}^{k} n_i$ where $P(w_i)$ has length $n_i$. The text solves the case $k = 3$ in $O(n \log n)$ time.
(a) Solve the minimum cover for $k = 2$ in linear time.
(b) Suppose $P(w_i) = (s_i, t_i)$ for each $i = 1, \ldots, k$, i.e., each keyword has just two positions. Give an $O(k \log k)$ algorithm to find the minimum cover $C$ for $w_1, \ldots, w_k$. HINT: suppose the minimal covers are $C_1, \ldots, C_m$ for some $m \geqslant 1$. Give an algorithm to list all the minimal covers. If $C_i = [c_i, d_i]$ and assuming $c_1 < c_2 < \cdots < c_m$, how do you find $C_1$? How do you find $C_{i+1}$ given $C_i$?
(c) Solve the general Google problem ($k$ is arbitrary and each word can have arbitrarily many occurrences in the file). HINT: if you used the hint from (b), it should be possible to generalize your solution. ◇

**Exercise 3.10:** Write a program to solve the Google multi-keyword for the case $k = 3$ as described in the text. Use your favorite programming language (C or Java without any Object-Oriented fanfare is recommended). Initially, assume $n$ is a power of 2. Indicate how to adapt your algorithm when $n$ is not a power of 2. ◇

**Exercise 3.11:** A cover $J = [s,t]$ for a multi-key search $w_1, \ldots, x_k$ was previously defined to mean that each $w_i$ occurs at least once in the positions of $J$. But suppose we modify this definition to mean that there is at least subsequence of positions $s \leqslant i_1 < i_2 < \cdots i_k \leqslant t$ such that $w_j$ occurs at position $i_j$ (for all $j$). Design an algorithm for the Keyword Cover Problem for this new definition of cover. ◇

**Exercise 3.12:** Consider the following problem: we are given an array $A[1..n]$ of numbers, possibly with duplicates. Let $f(x)$ be the number of times ("frequency") a number $x$ occurs. Given a number $k \geqslant 1$, we want to know whether there are $k$ distinct numbers $x_1, \ldots, x_k$ such that $\sum_{i=1}^{k} f(x_i) > n/2$. Call $\{x_1, \ldots, x_k\}$ a $k$-**majority set**.
(a) Solve this decision problem for $k = 1$.
(b) Solve this decision problem for $k = 2$.
(c) Instead of the previous decision problem, we consider the optimization version: find the smallest $k$ such that there are $k$ numbers $x_1, \ldots, x_k$ with $\sum_{i=1}^{k} f(x_i) > n/2$. ◇

End Exercises

## §4. Rote Method

They are "direct" as opposed to other transformation methods which we will introduce later. Although fairly straightforward, these direct methods may call for some creativity (educated guesses). We begin with the rote method, as it appears to require somewhat less guess work.

*"...at last, a method named after me!"* — Günter Rote (2010)

**¶10. What is rote?** The "rote method" refers to the idea of solving a recurrence by repeated expansion of a recurrence. Since such expansions can be done mechanically, this method has been characterized as rote.

Let us illustrate this method using the merge-sort recurrence (13): $T(n) = 2T(n/2) + n$. The important thing is that we can replace $n$ in this by any expression: plugging $n/2$ for $n$ in the recurrence, we get

$$T(n/2) = 2T(n/4) + n/2. \tag{22}$$

If we plug this back into the original recurrence, we get our second expansion in the following derivation:

$$\left.\begin{array}{rcll} T(n) & = & 2\boxed{T(n/2)}+n & \text{(first expansion)} \\ & = & 2\boxed{2T(n/4)+(n/2)}+n & \text{(second expansion, by (22))} \\ & = & 4\boxed{T(n/4)}+2n & \text{(simplify)} \\ & = & 4\boxed{2T(n/8)+(n/4)}+2n & \text{(third expansion)} \\ & = & 8T(n/8)+3n & \text{(simplify)} \end{array}\right\} \tag{23}$$

This is the expansion step. At this point, we may guess that the $i$th expansion, the formula is

$$(G)_i: \quad T(n) = 2^i T(n/2^i) + in. \tag{24}$$

To verify our guess, we use natural induction. Note that the formula (24) is true for $i = 1$ (it also holds for $i = 2$ and 3, but this is not logically necessary). We need an induction step: This amounts to expanding the formula once more:

$$\left.\begin{array}{rcll} T(n) & = & 2^i\boxed{T(n/2^i)}+in & \text{(guessed } i\text{th expansion)} \\ & = & 2^i\boxed{2T(n/2^{i+1})+n/2^i}+in & (i+1\text{st expansion)} \\ & = & 2^{i+1}T(n/2^{i+1})+(i+1)n, & \text{(simplify)} \end{array}\right\} \tag{25}$$

and noting that this confirms that the formula holds for $i+1$ (cf. formula $(G)_{i+1}$ in (24)).

Finally, we must choose a value of $i$ at which to stop this expansion. First consider the ideal situation where $n$ is a power of 2 and we choose $i = \lg n$. Then (24) yields $T(n) = 2^i T(n/2^i) + in = nT(1) + (\lg n)n$. Assume[6] that $T(1) = 1$, we obtain the solution $T(n) = n(1 + \lg n)$. This is a great solution, except for one problem: $i$ must be an integer, and it will not work when $n$ is not a power of 2. It makes no sense to pretend that $i$ is a real variable (as we did for $n$). In general, we may choose an integer close to $\lg n$: $\lceil \lg n \rceil$ or $\lfloor \lg n \rfloor$ will do. Let us choose

$$i = \lfloor \lg n \rfloor \tag{26}$$

as our stopping value. With this choice, we obtain $1 \leqslant n/2^i < 2$. We may exploit DIC to choose the initial condition

$$T(n) = n \lfloor \lg n \rfloor, \qquad \text{for } 0 < n < 2. \tag{27}$$

*Why not choose our usual $T(n) = 0$ for $0 < n < 2$?*

This yields the *exact* solution that for $n > 0$,

$$T(n) = n \lfloor \lg n \rfloor. \tag{28}$$

[6] In this example, since we have a concrete algorithm in view, we cannot invoke DIC arbitrarily: e.g., choosing "$T(1) = 0$" might work for some abstract recurrence, but does not make sense for the merging recurrence. In fact $T(1) = 1$ is obvious.

**¶11. Is it just rote?** To recap, there are four distinct stages in the rote method:

- **(E)** Expansion steps as in (23). This is the rote part. You can expand as many times as you like until you see the general pattern.
- **(G)** Guessing of a formula for the $i$th expansion, as in (24). This guess may require some creativity. Indeed, if we had not re-arranged the terms in our example in the suggestive manner, one might not see the pattern readily. So perhaps "rote" is a misnomer.
- **(V)** Verification of the formula as in (25). This step should be mechanical, and amounts to one more expansion step and re-arranging the terms into the desired form. One problem is that students sometimes do not do this step "honestly" (they jump to the expected conclusion).
- **(S)** Stopping criteria choice as in (26). You need to know when to stop expansion! Note you must choose $i$ to be a natural number. Thus, you cannot pick "$i = \lg n$" in (26), but need something like $i = \lceil \lg n \rceil$ or $i = \lfloor \lg n \rfloor$. According to DIC, you can pick any $i$ large enough that the recursive term $T(k)$ has an argument $k$ that is below some fixed constant (e.g., $k < 1$). Using DIC, you can declare $T(k)$ to be any value you like (usually $T(k) = 0$ is good).

  In general, your guess for the $i$-th expansion is in the form of a summation $\sum_{j=0}^{i-1} f(j)$ for some function $f$. If you stop at $m$-th expansion, you are left with the sum $\sum_{i=0}^{m-1} f(j)$. It just happens that for Mergesort, $f(i)$ is identically equal to $n$, and so the $\sum_{i=0}^{m-1} n$ is just $mn$ ($m = \lfloor \lg n \rfloor$). Unfortunately, we do not consider the open sum as adequate. Summation techniques will be taken up in its own section below. In any case, this fourth and last stage might be called the Stop-and-Sum stage.

Child's dilemma: *I can't spell* `banana` *because I don't know when to stop!*

Since the four stages are Expand, Guess, Verify and Stop-and-Sum, we also refer to the Rote Method as the **EGVS method**. When the method works, it can give you the exact solution. How can this method fail? It is clear that you can always perform expansions, but you may be stuck at the next step. For instance, try to expand the recurrence $T(n) = 2T(\lceil n/2 \rceil) + n$ in an exact form. The only way out is to give up on exact solution, and guess reasonable upper and/or lower bounds.

*Pronounce "EGVS" as "egg-us" (treat V as U like the Romans)*

**¶12. Exploiting DIC, Significance of DIC.** You may think of a recurrence as specifying an infinite family of problems: each problem corresponds to a choice of initial conditions. The nice part of DIC is that you get to choose your problem. Perhaps the main use of exploit DIC is to make your solution as simple as possible.

Let us illustrate this. In our rote solution of the merge-sort recurrence (13), we choose the initial condition: $T(n) = 0$ for $n < 2$ for its simplicity. But we ended up with the solution $T(n) = n \lfloor \lg n \rfloor$. This is admittedly simple, but the appearance of the floor function is a small annoyance. It also makes $T(n)$ discontinuous whenever $n$ is a power of 2.

Suppose that by DIC, we choose instead the following initial condition:

$$T(n) = n \lg n, \qquad (1 \leqslant n < 2).$$

It is a more "complicated" initial condition than before, but let us see the payoff. As before, after the $i$th expansion, we obtain

$$T(n) = 2^i T(n/2^i) + in, \quad (i \geqslant 1).$$

Plugging in $i = \lfloor \lg n \rfloor$, we obtain

$$\begin{aligned} T(n) &= 2^i T(n/2^i) + in \\ &= 2^i \left(\frac{n}{2^i} \lg\left(\frac{n}{2^i}\right)\right) + in \\ &= n(\lg n - i) + in \\ &= n \lg n \end{aligned}$$

*the "ultimate" in simplicity?*

for all $n \geqslant 1$. This solution is continuous and even simpler than (27).

We know of course that DIC is not realistic in the real world. The real world shows up in some relatively hard constraints on the initial conditions. How does it affect our solutions? They basically enforce some lower bounds on the implicit constants of our solution. For instance, if we want solutions of the form $T(n) = An\lg n + Bn$ for the Mergesort recurrence. Moreover, assume we have initial conditions of the form $C_0 \leqslant T(n) \leqslant C_1$ for $0 < n \leqslant n_1$. How will $A, B$ depend on $C_0, C_1, n_1$? See Exercises below (especially the "honest" Karatsuba or Mergesort Exercises).

Exercises

**Exercise 4.1:** No credit work: Rote is a discredited word in pedagogy, so we would like a more dignified name for this method. We could call this the "4-Fold Path". Suggest your own name for this method. In a humorous vein, what could EGVS (pronounced "egus") stand for? ♢

**Exercise 4.2:** Solve the following recurrence by the EGVS Method: $T(n) = 4T(n/2) + n^2$. ♢

**Exercise 4.3:** Use the EGVS Method to solve the following recurrences
(a) $T(n) = n + 8T(n/2)$.
(b) $T(n) = n + 16T(n/4)$.
(c) Can you generalize your results in (a) and (b) to recurrences of the form $T(n) = n + aT(n/b)$ when $a, b$ are positive powers of 2? ♢

**Exercise 4.4:** Solve the Karatsuba recurrence (21) using the Rote Method. ♢

**Exercise 4.5:** Give the exact solution for $T(n) = 2T(n/2) + n$ for $n \geqslant 1$ under the initial condition $T(n) = 0$ for $n < 1$. ♢

**Exercise 4.6:** Solve (19) assuming that $d(n) = n^\beta$ for some real $\beta$. NOTE: there will be three different cases, depending on the relationships between $\beta, a, b$. ♢

**Exercise 4.7:** (V. Shoup) You are given $n$ coins – they look identical, and all have the same weight except one, which is heavier than all the rest. You also have a balance scale, on which you can place one set of coins on one side, and another set of coins on the other, and the scale will tell you whether the two sets have the same weight, and if not, which is the heavier set. Suppose that performing one such weighing takes one minute, and in addition, you have to pay a fee of $m$ dollars, where $m$ is the total number of coins placed on the scale in that weighing. Design and analyze a strategy that will identify the heavy coin in $T(n) = O(\log n)$ minutes and at a cost of $C(n) = O(n)$ dollars. ♢

**Exercise 4.8:** Let us consider the following form of DIC, where we assume that

$$C_0 \leqslant T(n) \leqslant C_1$$

for $0 < n \leqslant n_1$, with the recurrence operative for $n > n_1$. Here, $C_0, C_1, n_1$ are positive constants. Give upper and lower bounds on the solution to the Mergesort recurrence $T(n) = 2T(n/2) + n$ in terms of $n_1, C_0, C_1$. NOTE: this question is interested in constant factors, so you must not hide them with asymptotic notations. ♢

**Exercise 4.9:** ("Honest Karatsuba") In this question, we want to take into account the multiplicative constants that are hidden by the Θ-notations. This is realistic or "honest".

(i) Argue that a more "honest" worst case recurrence for Karatsuba's algorithm should be

$$T(n) = 3T(\lceil n/2 \rceil + 1) + 5n + O(1). \tag{29}$$

Please justify all the constants $(1, 2, 3, 5)$ appearing in (29).

NOTE: since we are interested in constants in (29), we must tell you the cost to add two $n$-bit numbers: the cost is exactly $n$. Also, the cost to compute $Z$ from $Z_0, Z_1, Z_2$ is $2n$ (see ¶II.3, p. 7). But we don't really care about the $O(1)$ term in (29) (so we are still slightly abstract!).

(ii) Henceforth, assume $T(n)$ eventually satisfies the recurrence (29) *but without the* $O(1)$ *term*. We want to prove an upper bound $T(n)$ with explicit multiplicative constants. Consider a function of the form

$$U(n) = (n-3)^{\lg 3} - Kn \tag{30}$$

for some $K \geqslant 0$. Suppose $T(n) \leqslant U(n)$ (ev.). Determine the smallest possible value of the constant $K$. HINT: $\lceil n/2 \rceil + 1 \leqslant (n+3)/2$.

(iii) Argue why the upper bound (30) is STILL not "honest". How would do you suggest providing a realistic upper bound for $T(n)$? HINT: revoke DIC.

♢

**Exercise 4.10:** ("honest" Mergesort) This is analogous to the previous question. Suppose $T(n) = 2T(\lceil n/2 \rceil) + An$ where $A > 1$. Give an upper bound on $T(n)$ that takes into account the influence on $A$. ♢

End Exercises

## §5. Real Induction

The rote method, when it works, can give tight solutions to recurrences. Unfortunately, it does not work for most recurrences: while you can always expand, you may not be able to guess a simple and general formula for the $i$-th expansion. We now introduce a more widely applicable method, based on the idea of "real induction".

**¶13. Example of Real Induction.** You have probably never encountered real induction, so it is good to start with an example. Consider the recurrence

$$T(x) = T(x/2) + T(x/3) + x. \tag{31}$$

*Try using rote!!*

The student is encouraged to attempt the rote method on this recurrence, and see why it fails. Let us use real induction to prove an upper bound: suppose we guess that

$$T(x) \leqslant Kx(\text{ev.}) \text{ for some } K > 0. \tag{32}$$

Then we verify it "inductively":

$$\begin{array}{rcll} T(x) & = & T(x/2) + T(x/3) + x & \text{(by definition)} \\ & \leqslant & K\frac{x}{2} + K\frac{x}{3} + x & \text{(by ``inductive hypothesis'')} \\ & = & Kx\left(\frac{1}{2} + \frac{1}{3} + \frac{1}{K}\right) & \\ & \leqslant & Kx & \text{(provided } K \geqslant 6\text{)} \end{array}$$

Is this really a rigorous proof? We will justify this argument on general principles (see ¶17 below on growth functions).

To do this argument, we had to guess the upper bound $T(x) \leqslant Kx$. What if we had guessed $T(x) \leqslant Kx^2$? Well, your proof would have succeeded as well. In other words, this argument can confirm a particular guess. It says nothing about the optimality of the guess. But in reality, your proof yields hints on the tightness of the inequality.

We could likewise use real induction to confirm a guessed lower bound. The combined upper and lower bound can often lead to optimal bounds. In our example, suppose we guessed

$$T(x) \geqslant Kx(\text{ev.}) \text{ for some } K > 0. \tag{33}$$

Combined with (32), this would imply that $T(x) = \Theta(n)$. We just reverse the argument:

$$\begin{array}{rcll} T(x) & = & T(x/2) + T(x/3) + x & \text{(by definition)} \\ & \geqslant & K\frac{x}{2} + K\frac{x}{3} + x & \text{(by ``inductive hypothesis'')} \\ & = & Kx\left(\frac{1}{2} + \frac{1}{3} + \frac{1}{K}\right) & \\ & \geqslant & Kx & \text{(provided } K \leqslant 6\text{)} \end{array}$$

Unfortunately, the proviso that $K$ is not too large does not allow us to use the theorem in ¶17 below. We need[7] another argument.

## ¶14. Natural Induction.

Before introducing real induction, we first recall the well-known method of **natural induction**. This is a proof method based on induction over natural numbers. In brief, suppose $P(\cdot)$ is a natural number predicate, i.e., for each $n \in \mathbb{N}$, let $P(n)$ be a predicate (or proposition or assertion) about $n$. This predicate is either true or false for a given $n$. We simply write "$P(n)$" or, for emphasis, "$P(n)$ holds" when we want to assert "predicate $P(n)$ is true". The predicate $P(\cdot)$ is **valid** if it is true for all natural $n$:

$$(\forall n \in \mathbb{N})[P(n) \text{ holds}]. \tag{34}$$

For example, let $P_{10}(n)$ be the predicate

$$\text{``There is a prime number between } n \text{ and } n + 10 \text{ inclusive''.} \tag{35}$$

We may verify that $P_{10}(100)$ is true because 101 is prime, but $P_{10}(200)$ is false because 211 is the smallest prime larger than 200. Thus, $P_{10}(n)$ is invalid.

Next, consider the predicate defined as follows: $P_B(n) \equiv$"there is prime between $n$ and $2n - 1$", The validity of $P_B(\cdot)$ is called Bertrand's Postulate (1845). In 1850, Chebyshev proved this postulate. A common method to prove the validity of the predicate $P(\cdot)$ is "proof by natural induction". There are three steps in this proof:

(i) [*Natural Basis Step*] Show that there is some $n_1 \in \mathbb{N}$ that

$$(n \leqslant n_1) \Rightarrow P(n). \tag{36}$$

(ii) [*Natural Induction Step*] Show that if $n > n_1$ then

$$P(n-1) \Rightarrow P(n). \tag{37}$$

(iii) [*Principle of Natural Induction*] Invoke the principle of natural induction, which simply says that (i) and (ii) imply the validity of $P(\cdot)$, i.e., (34) holds.

Since step (iii) is independent of the predicate $P(\cdot)$, we only need to show the first two steps.

Example: suppose $F(n)$ satisfies the Fibonacci recurrence (1), with $F(0) = 0, F(1) = 1$. Define the predicate $P_F(n)$ that

$$\text{``}F(n) \leqslant \phi^n\text{''} \tag{38}$$

where $\phi = 1.618\cdots$ is the golden ratio. By definition, $\phi$ is the[8] positive number that satisfies the quadratic equation $x^2 - x - 1 = 0$ Clearly $P_F(0)$ and $P_F(1)$ holds since $F(0) = 0 < \phi^0$ and $F(1) = 1 < \Phi$.

[7] Even if we could prove (32) and (33) we must not conclude that $T(x) = 6x$ eventually because the initial conditions for the two proofs are chosen independently.

[8] There is another solution of the quadratic equation which is negative, denoted $\widehat{\phi}$.

In the proof by natural induction framework above, we can choose $n_0 = 1$. So the Natural Basis Step (i) is satisfied. It remains to show the Natural Induction Step (ii): we have for $n > n_0 = 1$,

$$\begin{array}{rcll} F(n) & = & F(n-1) + F(n-2) & \\ & < & \phi^{n-1} + \phi^{n-2} & \text{(Induction Hypothesis)} \\ & = & \phi^{n-2}(\phi+1) & \\ & = & \phi^{n-2}(\phi^2) & \text{(definition of } \phi) \\ & = & \phi^n & \text{(i.e., } P_F(n) \text{ holds)} \end{array}$$

Thus we have proved the validity of $P_F(n)$. Incidentally, we are not aware of any natural induction proof of Bertrand's Postulate.

An important variation of natural induction is the following: for any natural number predicate $P(\cdot)$, introduce a new predicate (the "star version of $P$") denoted $P^*(\cdot)$ defined as

$$P^*(n) : [m \leqslant n \Rightarrow P(m)]. \tag{39}$$

In the "Strong Natural Induction Step", we modify (37) in step (ii) into

$$P^*(n-1) \Rightarrow P(n). \tag{40}$$

It is easy to see that if we carry out the Natural Basis Step and the Strong Natural Induction Step, we have shown the validity of $P^*(n)$. Moreover, $P^*(\cdot)$ is valid iff $P(\cdot)$ is valid. Since $P^*(n)$ is stronger than $P(n)$, a proof of the validity of $P^*(\cdot)$ is called a **strong natural induction proof** of the validity of $P(\cdot)$.

**¶15. Real Induction.** We now introduce the real analogue of natural induction. For most students, this will be the first time they encounter the concept of "real induction". Indeed, it is rare in the mathematical literature. It arises naturally in some areas such as program verification [2], timed logic [13], and real computational models [4]. We believe it should also become an standard tool in the analysis of algorithms.

Real induction is applicable to **real predicates**, *i.e.*, a predicate $P(\cdot)$ such that for each $x \in \mathbb{R}$, we have a predicate denoted $P(x)$. For example, suppose $T(x)$ is a total complexity function that satisfies the Karatsuba recurrence (21) subject to the initial condition $T(x) = 1$ for $x \leqslant 10$. Let us define the real predicate

$$P(x) : [x \geqslant 100 \Rightarrow T(x) \leqslant x^2]. \tag{41}$$

As in (34), we want to prove the **validity** of the real predicate $P(\cdot)$, i.e.,

$$(\forall x \in \mathbb{R})[P(x) \text{ holds}]. \tag{42}$$

The continuous analogue of $P^*(n)$ in (39) is

$$P^*(x) : \Big[y \leqslant x \Rightarrow P(y)\Big] \tag{43}$$

We are ready to formulate the basic theorem of Real Induction:

**Theorem 1 (Principle of Real Induction)** *Let $P(x)$ be a real predicate. Suppose there exist real numbers $\delta > 0$ (gap constant) and $x_1$ (cutoff constant) such that*

(RB) [Real Basis Step] $P^*(x_1)$ *holds.*

(RI) [Real Induction Step] *(For all $x > x_1$)* $\big[P^*(x - \delta) \Rightarrow P(x)\big]$.

*Then $P(x)$ is valid, i.e., for all $x \in \mathbb{R}$, $P(x)$ holds.*

Before giving the proof, observe that this is truly the real analogue of strong natural induction: The cutoff constant $x_1$ corresponds to $n_1$ in Natural Basis (36); and the gap constant $\delta$ corresponds to the unit constant 1 of Natural Induction (37).

*Proof.* For $n \geqslant 1$, let

$$x_n := \begin{cases} x_1 & \text{if } n = 1, \\ x_{n-1} + \delta & \text{if } n \geqslant 2. \end{cases}$$

Also let $H_n$ be the half-line defined by $H_n := \{x \in \mathbb{R} : x \leqslant x_n\}$. The validity of $P(x)$ reduces to the validity of the natural number predicate

$$Q(n) : \left[ x \in H_n \text{ implies } P^*(x) \right]$$

By Real Basis Step (RB), we see that $x \in H_1$ implies $P(x)$ holds. Thus $Q(1)$ holds. We now prove that $Q(n)$ is valid by natural induction.[9] Suppose $n \geqslant 2$. To prove that $Q(n)$ holds, let $x \in H_n$. Then $x \leqslant x_n$ and $x - \delta \leqslant x_{n-1}$, i.e., $x - \delta \in H_{n-1}$. By induction hypothesis, $Q(n-1)$ holds, implying that $P^*(x - \delta)$ holds. By the Real Induction Step (RIH), $P^*(x-\delta)$ implies $P(x)$. Since this is true for all $x \in H_n$, we conclude that $Q(n)$ holds. By the Principal of Natural Induction, $Q(\cdot)$ is valid. This is clearly equivalent to the validity of $P(x)$. **Q.E.D.**

The above proof reduces Real Induction to Natural Induction. The principle behind this reduction is an intuitive property of real numbers: *Given $\delta > 0$ and real number $x$, there is a smallest integer $n(x)$ such that $x \leqslant n(x)\delta$.* E.g., Let $\delta = 0.2$. If $x = 19.9$ then $n(x) = 100$ since

$$19.8 = 99 \times 0.2 = 19.8 < x = 19.9 \leqslant 100 \times 0.2.$$

This property of real numbers is known as the **Archimedean Property**, after Archimedes of Syracuse (287–212 BC).

"*Give me a place to stand on [and lever long enough] and I can move the earth*"
– Archimedes
(quoted by Pappus of Alexandria)

**¶16. Initial conditions are messy! An Illustration** Let us apply the Principle of Real Induction for solving real recurrences. Its application requires the existence of two constants, $x_1$ and $\delta$, making it somewhat harder to use than natural induction.

Suppose $T(x)$ satisfies the recurrence

$$T(x) = x^5 + T(x/a) + T(x/b) \tag{44}$$

where $a \geqslant b > 1$ are real constants. Given $x_0 \geqslant 1$ and $K > 0$, let $P_{x_0,K}(x)$ be the following real predicate

$$P_{x_0,K}(x) \equiv \left[ x \geqslant x_0 \Rightarrow T(x) \leqslant Kx^5 \right]. \tag{45}$$

For simplicity, write $P(x)$ for $P_{x_0,K}(x)$ How do we establish the validity of $P(x)$?

**A priori relationships among $x_0, x_1$ and $\delta$:** Invariably, our real predicate $P(x)$ has the form

$$P(x) \equiv \left[ x \geqslant x_0 \Rightarrow Q(x) \right]$$

for some other predicate $Q(x)$, and some real value $x_0$. The example (45) above follows this pattern. In other words, $P(x)$ is the same as "$Q(x)$ (ev.)" except that we make $x_0$ explicit. This form is clearly suitable for proving properties such as domination between two complexity functions. This constant $x_0$ is needed to prove the Real Basis for $P(x)$. But the Real Basis introduces two other constants $x_1$ and $\delta > 0$. We must make sure that $x_1 \geqslant x_0 + \delta$ (at least). Otherwise, if $x_1 < x_0 + \delta$, then the fact that $P(x_1 - \delta)$ holds is a vacuous statement!

[9] The fact that we define $Q(n)$ only for $n \geqslant 1$ is clearly a trivial deviation from natural induction.

**Lemma 2** *Let $k_0 := a^{-5} + b^{-5}$ in the recurrence (44). If $k_0 < 1$ then for all $x_0 \geqslant 1$, there is a $K > 0$ such that $P_{x_0,K}(x)$ is valid.*

*Proof.* For any $x_0 \geqslant 1$, we will show that $P(x) = P_{x_0,K}(x)$ is valid for some $K > 0$. To apply the Principle of Real Induction, we must choose a cutoff constant $x_1$ and a gap constant $\delta > 0$. Pick any $x_1 > x_0$. Invoke Default Initial Condition to conclude that there is a $C > 0$ such that

*Choosing $x_1$*

$$T(x) \leqslant C$$

for all $x < x_1$. Now we must choose $K$ large enough, say

$$K \geqslant C/x_0^5. \tag{46}$$

Then for all $x \in [x_0, x_1)$, we have

$$\begin{array}{rcll} T(x) & \leqslant & C & \text{(by DIC)} \\ & \leqslant & Kx_0^5 & \text{(by (46))} \\ & \leqslant & Kx^5 & \text{(since } x \geqslant x_0 \geqslant 1\text{).} \end{array}$$

This establishes the Real Basis Step (RI) for $P(x)$ with cutoff $x_1$. To establish Real Induction Step (RI), we must now choose $\delta$. We further restrict the cutoff constant $x_1$ to satisfy

*Real Basis Step (RB) depends on the choice of $K$*

*choose gap $\delta$*

$$x_1 = ax_0. \tag{47}$$

*Further restriction on $x_1$*

Note that by assumption $a \geqslant b > 1$, and so $x_1 > x_0$. Thus for $x \geqslant x_1$, we have $x_0 \leqslant x/a \leqslant x/b$. Now we may choose the gap constant

$$\delta = x_1 - (x_1/b) = x_1 \frac{b-1}{b}. \tag{48}$$

This ensures that for $x \geqslant x_1$, we have $x/a \leqslant x/b = x - x\left(\frac{b-1}{b}\right) \leqslant x - \delta$. The Real Induction Hypothesis $P^*(x)$ says: for all $y \leqslant x$, $P(y)$ holds, i.e., $y \geqslant x_0 \Rightarrow P(y)$. We must show that $x \geqslant x_1$ and $P^*(x-\delta)$ implies $P(x)$:

*$y \in [x_0, x-\delta]$ is the sweet spot*

$$\begin{array}{rcll} T(x) & = & x^5 + T(x/a) + T(x/b) & (x \geqslant x_1) \\ & \leqslant & x^5 + K \cdot (x/a)^5 + K \cdot (x/b)^5 & \text{(by } P^*(x-\delta) \text{ and } x_0 \leqslant \frac{x}{a} \leqslant \frac{x}{b} \leqslant x - \delta\text{)} \quad (49) \\ & = & x^5(1 + K \cdot k_0) & (k_0 = a^{-5} + b^{-5}) \\ & \leqslant & Kx^5 & \quad (50) \end{array}$$

where the last inequality is guaranteed provided our choice of $K$ further[10] satisfies $1 + K \cdot k_0 \leqslant K$ or $K \geqslant 1/(1-k_0)$. This proves the Real Induction Step (RI). **Q.E.D.**

Notice that in this proof, we are doing a balancing act: in the recursive call to $T(x/a)$ and $T(x/b)$, we must ensure that both $x/a$ and $x/b$ lie in the range $[x_0, x-\delta]$. In short, we want $x/a$ to be at most $x - \delta$ but it cannot be too small (it would not work if $x/a$ is less than $x_0$). We achieve this by choosing $x_1$ to be sufficiently larger than $x_0$ (namely (47)). In a similar vein (Exercise), we can use real induction to prove a lower bound: there is a constant $k > 0$ such that $T(x) \geqslant kx^5$ (ev.). Therefore, we have shown $T(x) = \Theta(n^5)$ for the recurrence (44).

**¶17. On Partial Real Induction.** The previous example shows that the direct application of the Principle of Real Induction can be tedious, as we have to track constants such as $\delta, x_1$ and $K$. This tedium is mainly associated with justifying the Real Basis Step (RB); in contrast, the proof of the Real Induction Step (RI) is not tedious but seems instructive and natural. If you only prove (RI)

[10] So far, we have required $K \geqslant C/x_0^5$. So this is an additional constraint on $K$.

but not (RB), we say that you have[11] given a "partial real induction". The next theorem will provide conditions under which we can justify doing only "partial real induction".

We provide a bit more analysis of our main application in bounding complexity functions $T(x)$ by some function $F(x)$. Say we want to show $T(x) \leqslant F(x)$ (ev.). This translates to "$x \geqslant x_0 \Rightarrow T(x) \leqslant F(x)$." Call $x_0$ the **constant of eventuality**. This constant is distinct from the cutoff constant $x_1$. Indeed, we want $x_0 < x_1$. For $x \in [x_0, x_1]$, the bound "$T(x) \leqslant F(x)$" is asserted by invoking DIC. Moreover, the gap constant $\delta > 0$ should satisfy $\delta < x_1 - x_0$, so that the real induction step may be proved for $x \geqslant x_1$ by assuming that $y \in [x_0, x - \delta] \Rightarrow T(x) \leqslant F(x)$.

**¶18. Growth Functions and Justification of Default Real Basis.** The justification for "partial real induction" will depend on certain "growth" properties of functions.

A real function $f : \mathbb{R}^k \to \mathbb{R}$ is said to be a **growth function** if $f$ is eventually defined, eventually non-decreasing, positive, and unbounded in each of its variables. For instance, $f(x) = x^2 - 3x$ and $f(x, y) = x^y + x/\log x$ are growth functions, but $f(x) = -x$, $f(x) = 1 - 1/x$ and $f(x, y, z) = xy/z$ are not.

**Theorem 3** *We are given $f(x)$ and a real recurrence for $T(x)$:*

$$T(x) = G(x, T(g_1(x)), \ldots, T(g_k(x))).$$

*Under the hypothesis*

- *$G(x, t_1, \ldots, t_k)$, $f(x)$ and each $g_i(x)$ $(i = 1, \ldots, k)$ are growth functions.*
- *There is a constant $\delta > 0$ such that each $g_i(x) \leqslant x - \delta$ (ev. $x$),*

*we may conclude that*

$$T(x) \preceq f(x) \tag{51}$$

*provided we can choose some $K > 0$ large enough such that*

$$G(x, Kf(g_1(x)), \ldots, Kf(g_k(x))) \leqslant Kf(x)) \text{ (ev.}x). \tag{52}$$

We first clarify the import of this theorem: view (51) as a predicate $P(x)$. In the real induction proof for the validity of $P(x)$, the inequality (52) is really the Real Induction (RI) Step. This theorem says that under certain "growth assumptions", the Real Basis Step (RB) part is automatic.

Let us apply this theorem to our introductory recurrence (31). There we simply proved the (RI) part of the predicate $T(n) \leqslant Kx$ (ev.). The (RB) is automatic according to this theorem, because the functions $G(x, t_1, t_2) = x + t_1 + t_2$, $f(x) = x$, $g_1(x) = x/2$ and $g_2(x) = x/3$) are all growth functions.

*Proof.* Pick $x_0 > 0$ and $K > 0$ large enough so that all the "eventual premises" of the theorem are satisfied. In particular, $f(x), G(x, t_1, \ldots, t_k)$ and $g_i(x)$ are all defined, non-decreasing and positive when their arguments are $\geqslant x_0$. Also, $g_i(x) \leqslant x - \delta$ for each $i$ and $x \geqslant x_0$. Let $P(x)$ be the predicate

$$P(x) : \ x \geqslant x_0 \Rightarrow T(x) \leqslant Kf(x).$$

Pick

$$x_1 = \max\{g_i^{-1}(x_0) : i = 1, \ldots, k\}. \tag{53}$$

[11] This is similar to "partial correctness of algorithms". The correctness of an algorithm can be broken into two separation assertions: (i) the algorithm halts, and (ii), if the algorithm halts, then the output is correct. A "partial correctness proof" only shows assertion (ii).

The inverse $g_i^{-1}$ of $g_i$ is undefined at $x_0$ if there does not exist $y_i$ such that $g_i(y_i) = x_0$, or if there exists more than one such $y_i$. In this case, take $g_i^{-1}(x_0)$ in (53) to be any $y_i$ such that $g_i(y_i) \geqslant x_0$. We then conclude that for all $x \geqslant x_1$,

$$x_0 \leqslant g_i(x) \leqslant x - \delta.$$

By the Default Initial Condition (DIC), we conclude that for all $x \in [x_0, x_1]$, $P(x)$ holds. Thus, the Real Basis Step is verified. We now verify the Real Induction Step. Assume $x \geqslant x_1$ and $P^*(x-\delta)$. Then,

$$\begin{array}{rcll} T(x) & = & G(x, T(g_1(x)), \ldots, T(g_k(x))) & \\ & \leqslant & G(x, Kf(g_1(x)), \ldots, Kf(g_k(x))) & (\text{by } P^*(x-\delta)) \\ & \leqslant & Kf(x) & (\text{by } (52)). \end{array}$$

Thus $P(x)$ holds. By the Principle of Real Induction, $P(x)$ is valid. This implies $T(x) \preceq f(x)$. **Q.E.D.**

It is important to note that the theorem requires the ability to choose $K$ arbitrarily large. *This does not automatically hold for proving a lower bound, say $T(x) \succeq f(x)$.*

Let us see this theorem in action on the recurrence (44). We basically need to verify the following:

1. $f(x) = x^5$, $G(x, t_1, t_2) = x^5 + t_1 + t_2$, $g_1(x) = x/a$ and $g_2(x) = x/b$ are growth functions
2. $g_1(x) \leqslant x - 1$ and $g_2(x) \leqslant x - 1$ when $x$ is large enough (i.e., $\delta = 1$)
3. The inequality (52) holds if we choose $K \geqslant 1/(1 - k_0)$. This is just the derivation of (50) from (49).

The last step (52) is the most interesting, because it is most specific to the recurrence equation. From theorem 3 we conclude that $T(x) \preceq f(x)$.

It is clear that we can give an analogous theorem which can be used to easily establish lower bounds on $T(x)$. We leave this as an Exercise.

- One phenomenon that arises is that one often has to introduce a stronger induction hypothesis than the actual result aimed for. For instance, to prove that $T(x) = \mathcal{O}(x \log x)$, we may need to guess that $T(x) = Cx \log x + Dx$ for some $C, D > 0$. See the Exercises below.
- A real predicate $P$ can be identified with a subset $S_P$ of $\mathbb{R}$ comprising those $x$ such that $P(x)$ holds. The statement $P(x)$ can be generically viewed as asserting membership of $x$ in $S_P$, *viz.*, "$x \in S_P$". Then a principle of real induction is just one that gives necessary conditions for a set $S_P$ to be equal to $\mathbb{R}$. Similarly, a natural number predicate is just a subset of $\mathbb{N}$.

In the rest of this chapter, we indicate other systematic pathways; similar ideas are in lecture notes of Mishra and Siegel [14], the books of Knuth [11], Greene and Knuth [8]. See also Purdom and Brown [16] and the survey of Lueker [12].

### Exercises

**Exercise 5.1:** Consider the predicate (35).
(a) What is the smallest $n$ such that $P(n)$ is false?
(b) The Twin Prime Conjecture says that there are infinitely many $n$'s such that $n$ and $n+2$ are both prime. If the Twin Prime Conjecture is true, what can you say about the truthhood of $P(n)$? ♢

**Exercise 5.2:** Prove theorem 1, by reduction to natural induction. You can also use a proof by contradiction. ♢

**Exercise 5.3:** Consider the recurrence $T(x) = T(x/2) + T(x/3) + x$. In the text, we proved that $T(x) \leqslant Kx$ (ev.) for some $K > 0$. But suppose we had guessed (by the analogy to Mergesort recurrence) that $T(x) \leqslant Kx \lg x$ (ev.). Prove this by real induction. ♢

**Exercise 5.4:** Use real induction to provide good upper and lower bounds for the following:
(a1) $T(x) = T(x/2) + T(x/3) + T(x/4) + x$ HINT: Can you tell that $T(x)$ is superlinear? What is the watershed constant for this recurrence? (see §II.11).
(a2) $T(x) = T(x/2) + T(x/3) + T(x/4) + x^2$
(b1) $T(x) = 2T(x/2) + 3T(x/3) + x$
(b2) $T(x) = 2T(x/2) + 3T(x/3) + x^2$ ♢

**Exercise 5.5:** Suppose $T(x) = 5T(x/2) + x$. Show by real induction that $T(x) = \Theta(x^{\lg 5})$.
**HINT**: (1) Split this proof into separate upper and lower bound proofs.
(2) We recommend that you first prove Partial Real Induction (i.e., by assuming Real Basis holds). But you need to show the Real Basis later. Remember that you need to introduce explicit constants like $K > 0$, as in $T(x) \leqslant Kx^{\lg 5}$ (ev.).
(3) For the upper bound, you may need to strengthen the induction hypothesis to something like $T(x) \leqslant Kx^{\lg 5} - g(n)$ (ev.). for some $g(n) > 0$. ♢

**Exercise 5.6:** In the previous question, we have a recurrence whose solution is $T(n) = \Theta(n^c)$ for some $c > 0$. When proving the *upper bound* by real induction, we are obliged to strengthen the obvious induction hypothesis. In this question, we ask you to state a similar recurrence with solution $T(n) = \Theta(n^c)$ but when prove the *lower bound* we must strengthen the obvious hypothesis. ♢

**Exercise 5.7:** We extend one of the previous problems: consider the recurrence $T(x) = 5T(x/2) + x^c$ where $c$ is a constant. Under what condition on $c$ is $T(x) = \Theta(x^{\lg 5})$? ♢

**Exercise 5.8:** Given $T(x) = 9T(x/2) + x^3$, whow by real induction that $T(x) \leqslant K9^{\lg x} - K'x^3$. What is the smallest value of $K'$ you can use? ♢

**Exercise 5.9:** Consider equation (13), $T(n) = 2T(n/2) + n$. Fix any $k > 1$. Show by induction that $T(n) = \mathcal{O}(n^k)$. Which part of your argument suggests to you that this solution is not tight? ♢

**Exercise 5.10:** Consider the recurrence $T(n) = n + 10T(n/3)$. Suppose we want to show $T(n) = \mathcal{O}(n^3)$.
(a) Give a proof by real induction.
(b) Suppose $T(n) = n + 10T((n + K)/2)$ for some constant $K$. How does your proof in (b) change? ♢

**Exercise 5.11:** Let $T(n) = 2T(\frac{n}{2} + c) + n$ for some $c > 0$.
(a) By choosing suitable initial conditions, prove the following bounds on $T(n)$ by induction, and *not* by any other method:
(a.1) $T(n) \leqslant D(n - 2c)\lg(n - 2c)$ for some $D > 1$. Is there a smallest $D$ that depends only on $c$? Explain. Similarly, show $T(n) \geqslant D'(n - 2c)\lg(n - 2c)$ for some $D' > 0$.
(a.2) $T(n) = n \lg n - o(n)$.
(a.3) $T(n) = n \lg n + \Theta(n)$.
(b) Obtain the exact solution to $T(n)$.
(c) Use your solution to (b) to explain your answers to (a). ♢

**Exercise 5.12:** Generalize our principle of real induction so that the constant $\delta$ is replaced by a real function $\delta : \mathbb{R} \to \mathbb{R}_{>0}$. What additional assumptions do we need? ♢

**Exercise 5.13:** (Gilles Dowek, "Preliminary Investigations on Induction over Real Numbers", manuscript 2002).

(a) A set $S \subseteq \mathbb{R}$ is closed if every limit point of $S$ belongs to $S$. Let $P(x)$ be a real predicate $P(x)$. Assume $\{x \in \mathbb{R} : P(x) \text{holds}\}$ is a closed set. Suppose

$$P(a). \wedge .(\forall c \geqslant a)[P(c). \Rightarrow .(\exists \varepsilon)(\forall y)[c \leqslant y \leqslant c + \varepsilon \Rightarrow P(y)]]$$

Conclude that $(\forall x \geqslant a)P(x)$.

(b) Let $a, b \in \mathbb{R}$ and $\alpha, \beta : \mathbb{R} \to \mathbb{R}$ such that for all $x$, $\alpha(x) \geqslant 0$ and $\alpha(x) > 0$. Suppose $f$ is a differentiable function satisfying

$$f(a) = bf'(x) = -\alpha(x)f(x) + \beta(x)$$

then for all $x \geqslant a$, $f(x) > 0$. Intuition: If $f(x)$ is the height of an object at time $x$, then the object will never reach the ground, *i.e.*, $f(x) > 0$. ◇

End Exercises

## §6. Basic Sums

In this section, we discuss some well-known basic sums and their role in solving recurrences.

**¶19. Rote expansion of the Master Recurrence.** As motivation, let us return to the rote or EGVS method. We have used it for the Mergesort recurrence (13). Let us try apply the technique to the general Master Recurrence (19) which is

$$T(n) = aT(n/b) + f(n)$$

for $a > 0$ and $b > 1$. Expanding, guessing and verifying yields:

$$\begin{aligned} T(n) &= a\boxed{T(n/b)} + f(n) \\ &= a^2\boxed{T(n/b^2)} + af(n/b) + f(n) \\ &= \cdots \\ &= a^i\boxed{T(n/b^i)} + \sum_{j=0}^{i-1} a^j f(n/b^j). \end{aligned}$$

Let us stop when $i = \lfloor \log_b n \rfloor$. Then $n/b^i < b$. We may assume DIC with $T(n) = 0$ for $n < b$. This gives us

$$T(n) = \sum_{j=0}^{\lfloor \log_b n \rfloor} a^j f(n/b^j). \tag{54}$$

Upon stopping, unlike in the Mergesort case, we now get an **open sum**, i.e., a sum with an unbounded number of summands depending on $n$. We do not regard an open sum as a satisfactory solution. We must convert this into a closed form. This conversion is the topic of this section.

**¶20. The Standard Recurrence and Descending Sums.** Basically, the EGVS method has transformed the Master Recurrence into a recurrence of the form

$$\boxed{T(n) = T(n-1) + f(n). \qquad (55)}$$

We shall call this the **standard recurrence**. Our goal in the following sections is to show systematic ways to reduce many recurrences into this standard form. Trivially, (55) has the following open sum as solution

$$T(n) = \sum_{i=1}^{n} f(i), \tag{56}$$

assuming $T(0) = 0$ and $n$ is integer.

In the solution (56) we have assumed that $n$ is integer. We shall see examples later where $n$ is not necessarily integer. Let us introduce some general notations that befit our intention of "going totally real". For any real numbers $a, b$, we define two kinds of sums of $f$-values over this real interval $[a, b]$:

*$\sum_{x\geqslant 1}^{\pi} x = 3\pi - 3$, but $\sum_{x=1}^{\pi} x = 6$ where $\pi = 3.1415\ldots$.*

$$\left.\begin{array}{lcll} \sum_{i=a}^{b} f(i) & := & f(a) + f(a+1) + f(a+2) + \cdots + f(a + \lfloor b-a \rfloor) & (ascending) \\ \sum_{i\geqslant a}^{b} f(i) & := & f(b) + f(b-1) + f(b-2) + \cdots + f(b - \lfloor b-a \rfloor) & (descending) \end{array}\right\} \tag{57}$$

We call them **ascending** and **descending** $f$**-summations**. Note that the last term in the ascending sum is $f(a + \lfloor b-a \rfloor)$, which is not necessarily equal to $f(b)$. Likewise, the last term in the descending sum is $f(b - \lfloor b-a \rfloor)$, which is not necessarily equal to $f(a)$. Moreover, the both sums are empty iff $b < a$. The value of empty sums is defined to be 0.

The difference between these two notations lies in a minute detail – in the way we express the lower bound on the index $i$ of the summation: "$\sum_{i\geqslant a}^{b}$" versus "$\sum_{i=a}^{b}$".

*pay close attention to this detail!*

Descending sums are more natural in solving recurrences but in other settings ascending sums seem more natural. Here is a simple transformation between ascending and descending sums:

$$\boxed{\sum_{i\geqslant a}^{b} f(i) = \sum_{i=0}^{b-a} f(b-i). \qquad (58)}$$

The right-hand side is also equal to $\sum_{i=0}^{\lfloor b-a \rfloor} f(b-i)$. Even when $f(x)$ is a partial function, these sums are well-defined using the convention that *undefined summands are replaced by* 0. In recognition of our interest in descending sums, we introduce a convenient notation: for any complexity function $f$, let

*convention for summing over partial functions*

$$S_f(n) := \sum_{i\geqslant 1}^{n} f(i). \tag{59}$$

and thus the solution to our standard recurrence (55) is

$$T(n) = S_f(n). \tag{60}$$

**¶21. What Does It Mean to Solve a Recurrence?** If the open sum in the RHS of (56) is unsatisfactory, what is satisfactory? Let us get a hint using a simple example. Suppose $f(n) = n$ in (56). Then we know how to convert the open sum into a **closed sum**:

$$T(n) = \sum_{i=1}^{n} f(i) = \sum_{i=1}^{n} i = \binom{n+1}{2} = \frac{n(n+1)}{2} = \Theta(n^2).$$

Indeed, we would be perfectly happy with the answer "$T(n) = \Theta(n^2)$" even though the answer is really $\binom{n+1}{2}$ — remember that we are generally interested in $\Theta$-order answers in this book. The reason we are happy with the answer $\Theta(n^2)$ is because $n^2$ is a "familiar function". So this section is about how we can write some "basic sums" in terms of familiar functions. These sums are the ones you must have under your belt.

*I see! "Solving" means to relate to known functions*

**¶22. On Familiar Functions.** So we conclude that "solving a recurrence" means expressing the recurrence function in the Θ-order of a suitable class of "familiar functions" such as $n^2$ or $n \log n$. More precisely, we may define the class of familiar functions to comprise certain "simple functions" and is closed under the following "familiar" operations:

| | |
|---|---|
| sum | $f + g$ |
| product | $fg$ |
| logarithm | $\log f$ |
| exponentiation | $f^g$ |
| functional composition | $f \circ g$ |

The "simple functions" may be taken to be either the identity $f(n) \equiv n$, or the constants functions $f(n) \equiv c$ ($c \in \mathbb{R}$). Thus, familiar operations include polynomials $f(n) = n^k$, iterated logarithms $f(n) = \log^{(k)} n$, simple exponentials $f(n) = c^n$ ($c > 0$).

It is useful to extend the familiar functions by allowing somewhat less obvious operations: taking factorials $n!$, binomial coefficients $\binom{n}{k}$, and harmonic numbers $H_n$ (see below). Since these operations can be tightly bounded by other familiar operations, they may be considered familiar in an extended sense. In addition to the above functions, two very slow growing functions arise naturally in algorithmic analysis. These are the log-star function $\log^* x$ (see Appendix) and the inverse Ackermann function $\alpha(n)$ (see Chapter XII). We will consider them familiar, although functional compositions with these strange functions are only "familiar" in our rather technical sense!

We refer the reader to Appendix A in this lecture for basic properties of the exponential and logarithm function. A useful relation is the following:

**Lemma 4** *For all real $0 < a < b$, real $c > 1$, and integer $k \geqslant 1$, we have:*

$$1 \lll \lg^{(k+1)} n \lll \lg^{(k)} n \lll n^a \lll n^b \lll c^n$$

What follows is a brief introduction to some common familiar functions, *expressed as solutions to summations.* The vast majority of the summations in this book can be reduced to one of these.

**¶23. Generalized arithmetic series.** The term **series** in analysis refers to an infinite summation $S = \sum_{i=0}^{\infty} a_i$ where each $a_i$ is a number (or some summable objects). The fundamental question about series is whether it gives a meaningful value. This is reduced to questions about its $n$-th **partial sum** $S_n = \sum_{i=1}^{n} a_i$. One of the simplest example is the well-known **arithmetic series**,

$$S_n \quad := \quad 1 + 2 + \cdots + n - 1 + n = \sum_{i=1}^{n} i. \tag{61}$$

Here is the usual one-line argument that $S_n = \binom{n+1}{2}$:

$$2S_n = \Big(\sum_{i=1}^{n} i\Big) + \Big(\sum_{i=1}^{n} (n+1-i)\Big) = \sum_{i=1}^{n} (n+1) = n(n+1).$$

This is the algebraic form of the well-known "proof by picture" where you draw two congruent staircases, each representing the desired sum; we can put these two (Lego-block) staircases together to form a rectangle of area $2S_n = n(n+1)$.

*Egg Drop Problem (from a job interview).* The Empire State Building in Manhattan has 102 floors.

Assume that there is a unique $1 \leqslant x^* \leqslant 102$ such that if you can drop an (artificial) egg from any floor $\geqslant x^*$, the egg will break. But if you drop it from any floor $\leqslant x^* - 1$, the egg will not break. Suppose you have a budget of $k \geqslant 1$ eggs to discover the value of $x^*$. Let $D(k)$ be the (worst case) number of "egg drops" necessary to discover the value of $x$ within the budget. Clearly $D(1) = x^*$: you drop at floors $1, 2, \ldots,$ until you reach a floor when the egg broke. In the worst case, $x^* = 102$. What is $D(2)$? Answer: $D(2) = 14$. Explanation at the end of this chapter.

More generally, for fixed $k \geqslant 1$, we have the "arithmetic series of order $k$" $S^k := \sum_{i=0}^{\infty} i^k$. Its $n$-th partial sum is given by:

$$S^k_n := \sum_{i=1}^{n} i^k = \Theta(n^{k+1}). \tag{62}$$

In proof, we have

$$n^{k+1} > S^k_n > \sum_{i=\lceil n/2 \rceil}^{n} (n/2)^k \geqslant (n/2)^{k+1}.$$

We can get slightly more precise for $S^k_n$ using integrals,

$$\frac{n^{k+1}}{k+1} = \int_0^n x^k dx < S^k_n < \int_1^{n+1} x^k dx = \frac{(n+1)^{k+1} - 1}{k+1},$$

yielding

$$S^k_n = \frac{n^{k+1}}{k+1} + \mathcal{O}_k(n^k). \tag{63}$$

Observer that the preceding derivation does not require $k$ to be integer: the derivation is valid for any real $k > -1$. When $k$ is integer, exact formulas for $S^k_n$ can be systematically derived (Exercise in next section). For instance, $S^2_n = \frac{n(n+1)(2n+1)}{6}$ and $S^3_n = (S_n)^2$ (verify this by induction).

*Don't worry about the integrals here: to achieve* $\Theta$*-bounds, we are always able to replace calculus by elementary methods.*

**¶24. Geometric series.** The **geometric series** is $S(x) := \sum_{i=0}^{\infty} x^i$ for some number $x$. We assume $x$ is real. Its $n$th partial sum is given by

$$\begin{aligned} S_n(x) &:= \sum_{i=0}^{n-1} x^i \\ &= \begin{cases} \frac{x^n - 1}{x - 1} & \text{if } x \neq 1 \\ n & \text{if} x = 1. \end{cases} \end{aligned} \tag{64}$$

In proof, note that $xS_n(x) - S_n(x) = x^n - 1$. Next, letting $n \to \infty$, we get the series

$$\begin{aligned} S_\infty(x) &:= \sum_{i=0}^{\infty} x^i \\ &= \begin{cases} \infty & \text{if } x \geqslant 1 \\ \uparrow \text{ (undefined)} & \text{if } x \leqslant -1 \\ \frac{1}{1-x} & \text{if } |x| < 1. \end{cases} \end{aligned}$$

Why is $S_\infty(-1)$ (say) considered undefined? For instance, writing

$$\begin{aligned} S_\infty(-1) &= 1 - 1 + 1 - 1 + 1 - 1 + \cdots \\ &= (1-1) + (1-1) + (1-1) + \cdots \\ &= 0 + 0 + 0 + \cdots, \end{aligned}$$

we conclude $S_\infty(-1) = 0$. But writing

$$\begin{aligned} S_\infty(-1) &= 1 - 1 + 1 - 1 + 1 - \cdots \\ &= 1 - (1-1) + (1-1) - \cdots \\ &= 1 + 0 + 0 + \cdots, \end{aligned}$$

we conclude $S_\infty(-1) = 1$. So that we must consider this sum as having no definite value, i.e., undefined. Again,

$$\begin{aligned} S_\infty(-1) &= 1 - 1 + 1 - 1 + 1 - \cdots \\ &= 1 - S_\infty(-1), \end{aligned}$$

and we conclude that $S_\infty(-1) = 1/2$. In fact, $S_\infty(-1)$ can take infinitely many possible values in this way. This provides a strong case why $S_\infty(-1)$ should be regarded as undefined.

what 19th century mathematicians learned: *handle infinite sums with great care*

Viewing $x$ as a formal[12] variable, the simplest infinite series is $S_\infty(x) = \sum_{i=0}^{\infty} x^i$. It has a very simple closed form solution,

$$\sum_{i=0}^{\infty} x^i = \frac{1}{1-x}. \tag{65}$$

*The one infinite series to know!*

Viewed numerically, we may regard this solution as a special case of (64) when $n \to \infty$; but avoiding numerical arguments, it can be directly derived from the formal identity $S_\infty(x) = 1 + xS_\infty(x)$. We call $\sum_{i=0}^{\infty} x^i$ the **mother of series** because[13] from the formal solution to this series, we can derive solutions for many related series, including finite series. In fact, for $|x| < 1$, we can derive equation (64) by plugging equation (65) into

$$S_n(x) = S_\infty(x) - x^n S_\infty(x) = (1 - x^n)S_\infty(x).$$

By differentiating both sides of the mother series with respect to $x$, we get:

$$\begin{aligned} \frac{1}{(1-x)^2} &= \sum_{i=1}^{\infty} i x^{i-1} \\ \frac{x}{(1-x)^2} &= \sum_{i=1}^{\infty} i x^i \end{aligned} \tag{66}$$

This process can be repeated to yield formulas for $\sum_{i=0}^{\infty} i^k x^i$, for any integer $k \geqslant 2$. Differentiating both sides of equation (64), we obtain the finite summation analogue:

$$\begin{aligned} \sum_{i=1}^{n-1} i x^{i-1} &= \frac{(n-1)x^n - nx^{n-1} + 1}{(x-1)^2}, \\ \sum_{i=1}^{n-1} i x^i &= \frac{(n-1)x^{n+1} - nx^n + x}{(x-1)^2}, \end{aligned} \tag{67}$$

$$\tag{68}$$

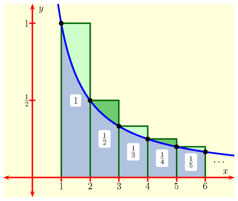
Combining the infinite and finite summation formulas, equations (66)–(67), we also obtain

$$\sum_{i=n}^{\infty} i x^i = \frac{nx^n - (n-1)x^{n+1}}{(1-x)^2}. \tag{69}$$

We may verify by induction that these formulas actually hold for all $x \neq 1$ when the series are finite. In general, for any $k \geqslant 0$, we obtain formulas for the **geometric series of order** $k$:

$$\sum_{i=1}^{n-1} i^k x^i. \tag{70}$$

The infinite series have finite values only when $|x| < 1$.



## ¶25. Harmonic series.

For natural numbers $n \geqslant 1$, the $n$th **harmonic number** is defined as

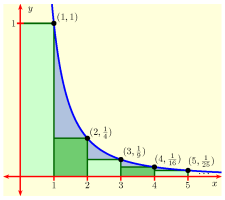
$$H_n := 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}. \tag{71}$$

[12] I.e., as an uninterpreted symbol rather than as a numerical value. Thereby, we avoid questions about the sum converging to some unique numerical value.

[13] This terminology arose in 1990, during the Gulf War when Saddam Hussein declared the "mother" of all battles. Suddenly, many things are declared the "mother of ..." and this appellation seems to fit $S_\infty(x)$.

We can give easy estimates of $H_n$ using calculus (see margin):

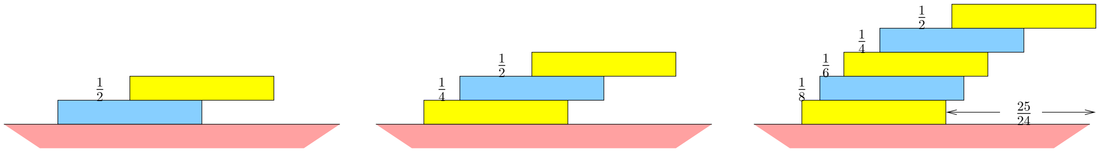
$$H_n < 1 + \int_1^n \frac{dx}{x} < 1 + H_n.$$



But $\int_1^n \frac{dx}{x} = \ln n$. This proves that

$$H_n = \ln n + g(n), \qquad \text{where } 0 < g(n) < 1. \tag{72}$$

Note that ln is the natural logarithm (appendix A).

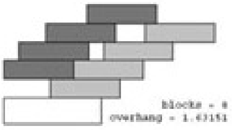


(a) $n = 2$ (b) $n = 3$ (c) $n = 5$

Figure 4: Stacking bricks with maximum overhang: for $n = 5$, overhang exceeds one brick length!

*Does your architecture friend know this?*

Harmonic numbers arise naturally in the analysis of algorithms. But here is a "physical" application of harmonic numbers: Suppose you have a set of $n \geqslant 2$ bricks. The bricks are identical and have unit length. We want to stack the bricks so that the overhang is as large as possible. For instance, if $n = 2$, the overhang is 1/2 since we can put one brick over the other such that the center of gravity of the top brick is above the edge of the bottom brick. This is illustrated in Figure 4(a). The case of $n = 3$, we may check that the overhang is 3/4 (Figure 4(b)). An obvious question is whether we can make the overhang arbitrarily large (provided $n$ is large enough)? Somewhat surprisingly, the answer is 'yes'. See Figure 4(c) for the case $n = 5$: in this case, the overhang is 25/24, already exceeding the length of a single brick! How many bricks do we need to have an overhang exceeding two brick lengths? In general, the overhang is $\frac{1}{2}H_{n-1}$ (Exercise). As $H_n$ is about $\ln n$, the overhang goes to infinity (albeit very slowly) as $n \to \infty$.
For more information, see the fascinating book "How Round is Your Circle? Where Engineering and Mathematics Meet", by John Bryant and Chris Sangwin (Princeton University Press, 2008). This solution is based on an assumption that you stack at most one brick on another. What if you allow more than one? You can do a lot better than the above classical solution! Mike Paterson and Uri Zwick (2009, American Math. Monthly) have investigated the case of multiple stacking. The maximum overhang for 8 bricks are illustrated in the margin here.



We can view (72) as a special case of our descending sums $S_f(n)$ where $f(n) = 1/n$. Then for all real $n$, $H_n = S_f(n) = \sum_{i \geqslant 1}^{n} \frac{1}{i}$. Here is a more precise estimate for $g(n)$: for $n \geqslant 1$,

$$\gamma + \frac{1}{2n} - \frac{1}{8n^2} < g(n) < \gamma + \frac{1}{2n} \tag{73}$$

where $\gamma = 0.577...$ is **Euler's constant**. See Polya and Szego, Problems and Theorems in Analysis, Volume I, Springer-Verlag, Berlin (1972).

We can deduce asymptotic properties of $H_n$ without calculus: if $n = 2^N$ (for some $N \geqslant 1$), then the terms in the defining summation of $H_n$ can be put into $N$ groups as follows

$$H_n = \sum_1 + \sum_2 + \cdots + \sum_N + \frac{1}{n} \tag{74}$$

where the $k$th group $\sum_k$ is defined as $\sum_{i=2^{k-1}}^{2^k-1}\frac{1}{i}$. Notice that the last term $1/n$ is not in any group. For example

$$H_8 = \underbrace{\frac{1}{1}}_{\Sigma_1} + \underbrace{\frac{1}{2}+\frac{1}{3}}_{\Sigma_2} + \underbrace{\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}}_{\Sigma_3} + \frac{1}{8}.$$

Since $\sum_k$ has $2^{k-1}$ terms, and each term is between $1/2^k$ and $1/2^{k-1}$, we obtain

$$\begin{aligned} 2^{k-1}\frac{1}{2^k} \leqslant \sum_k \leqslant 2^{k-1}\frac{1}{2^{k-1}} \\ 1/2 \leqslant \sum_k \leqslant 1. \end{aligned} \tag{75}$$

This proves[14] that

$$\begin{aligned} \tfrac{1}{2}N \leqslant H_n \leqslant N + \frac{1}{n} \\ \tfrac{1}{2}\lg n \leqslant H_n \leqslant \lg n + \frac{1}{n} \end{aligned} \tag{76}$$

when $n$ is a power of 2. Extrapolating to all values of $n$, we obtain

$$\tfrac{1}{2}\lfloor \lg n \rfloor \leqslant H_n \leqslant \lceil \lg n \rceil + \frac{1}{n}$$

Since we may choose $N$ as big[15] as we like, we have proved the following:

**Lemma 5**
*(a)* $H_n = \Theta(\lg n)$.
*(b)* $\lg n$ *is eventually unbounded, i.e.,* $\lg(n) \gg 1$.

The technique in this demonstration is again used to prove Theorem 9, and fully developed in [19]. In the next section, we will generalize $H_n$ to $H^{(\alpha)}(n)$ for real numbers $\alpha, n$.

**¶26. Stirling's Approximation.** So far, we have treated open sums. If we have an open product such as the factorial function $n!$, we can convert it into an open sum by taking logarithms. This method of estimating an open product may not give as tight a bound as we wish (why?). For the factorial function, there is a family of more direct and sharp bounds that are collectively called **Stirling's approximation**. The following Stirling approximation is from Robbins (1955) and it may be committed to memory:

$$n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n}\, e^{\alpha_n}$$

where

$$\frac{1}{12n+1} < \alpha_n < \frac{1}{12n}.$$

Sometimes, the bound $\alpha_n > (12n)^{-1} - (360n^3)^{-1}$ is useful [5]. Up to $\Theta$-order, Stirling's approximation simplifies to

$$n! = \Theta\left(\left(\frac{n}{e}\right)^{n+\frac{1}{2}}\right).$$

[14] For $N \geqslant 3$, the term $1/n$ could be ignored because we can count it as part of $\Sigma_2$. Note that $\frac{1}{2} \leqslant \Sigma_2 \leqslant 1$ still hold true after absorbing this extra term.

[15] A function $f$ might be unbounded, but not in the eventual sense. E.g., $f(x) = 1/(1-x)^2$.

**¶27. Binomial theorem.** The familiar (finite) form of the binomial theorem says: for any real $x$ and natural number $n$,

$$(1+x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i = 1 + nx + \frac{n(n-1)}{2}x^2 + \cdots + x^n. \tag{77}$$

In solving real recurrences, it is useful to generalize this theorem to $(1+x)^p$ for any real number $p$. In general, the binomial function $\binom{n}{i}$ may be extended to all real $p$ and integer $i$ as follows:

$$\binom{p}{i} = \begin{cases} 0 & \text{if } i < 0 \\ 1 & \text{if } i = 0 \\ \frac{p(p-1)\cdots(p-i+1)}{i(i-1)\cdots 2\cdots 1} & \text{if } i > 0. \end{cases}$$

We use Taylor's expansion for a function $f(x)$ at $x = a$:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \cdots$$

where $f^{(n)}(x) = \frac{d^k n}{df^k}x$. This expansion is defined provided all derivatives of $f$ exist and the series converges. Applied to $f(x) = (1+x)^p$ for any real $p$ at $x = 0$, we get the desired binomial theorem for real exponents:

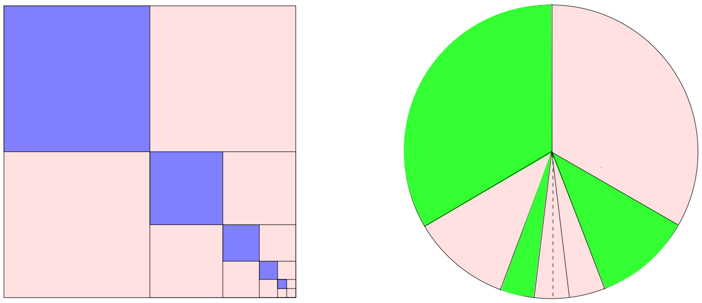
$$\begin{aligned} (1+x)^p &= 1 + px + \frac{p(p-1)}{2!}x^2 + \frac{p(p-1)(p-2)}{3!}x^3 + \cdots \\ &= \sum_{i \geqslant 0} \binom{p}{i} x^i. \end{aligned}$$

See Knuth [11, p. 56] for Abel's generalization of the binomial theorem.

### Exercises

**Exercise 6.1:** Show Lemma 4. For logarithms, please use direct inequalities (no calculus). ◇

**Exercise 6.2:** The Mother of Series is very important, and you should recognize it in its many forms. For this problem, you must not directly use the formula for the geometric series.
(a) Proof-by-Picture. Let $S_4 = \frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \frac{1}{256} + \frac{1}{1024} + \cdots = \sum_{i=1}^{\infty}(1/4)^i$. Use Figure 5(a) to determine the value of $S_4$.
(b) Let $S_3 = \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \frac{1}{81} + \cdots = \sum_{i=1}^{\infty}(1/3)^i$. Again, use Figure 5(b) to determine the value of $S_3$.
(c) Generalize the arguments of (a) and (b) to $S_k = \sum_{i=1}^{\infty} k^{-i}$. ◇



(a) $S_4$ (b) $S_3$

Figure 5: Proof-by-Picture

**Exercise 6.3:** For all natural number $n$, $(1+2+3+\cdots+n)^2 = 1^3+2^3+3^3+\cdots+n^3$.
(a) Give an inductive proof.
(b) Proof-by-Picture. Hint: Figure 6. $\diamondsuit$

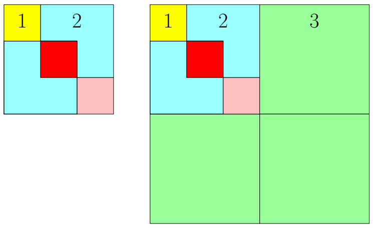


Figure 6: Cubes from squares: $n = 2, 3$

**Exercise 6.4:** Let $n = 2^N$ for $N \geqslant 1$. Sharpen (76) to $1 + (N/2) \leqslant H_n \leqslant N + \frac{1}{n}$. HINT: break $H_n$ into $N$ sums of the form $\Sigma_k = \sum_{i=2^{k-1}+1}^{2^k} \frac{1}{i}$. $\diamondsuit$

**Exercise 6.5:** The solution of the Master Recurrence in (54) is in ascending form: $T(n) = \sum_{j=0}^{\log_b n} a^j f(n/b^j)$. Use (58) to transform it into the descending sum. How is your answer related to the descending sum $S_g(\log_b n) = \sum_{i \geqslant 1}^{\log_b n} g(i)$ where $g(i) := f(b^i)/a^i$? $\diamondsuit$

**Exercise 6.6:** Let $S(27)$ denote the minimal height of a tree program to sort 27 elements (Chapter I ¶19).
(a) What is $S(27)$ and describe how you would go about computing this number, armed with only a pocket calculator.
(b) Describe any pitfalls and issues with numerical errors in this computation, and how you avoided them. $\diamondsuit$

**Exercise 6.7:** Strengthen the lower bounds in Lemma 4 from $\neq \Omega(f(n))$ to $= o(f(n))$. $\diamondsuit$

**Exercise 6.8:** Let $h(n)$ denote the maximum overhang for a stable stack of $n$ bricks. Clearly $h(1) = 0$ and $h(2) = 1/2$.
(a) Prove that $h(n) = \sum_{i=1}^{n-1} \frac{1}{2i} = \frac{1}{2} H_{n-1}$.
(b) Use a calculator to calculate the smallest $n$ such that $h(n) > 100$. Is such a stack realistic?
HINT: An ***n*-stack** of $n$ bricks where the first brick is at the bottom, and the $i+1$st brick is placed on top of the $i$th brick (for $i = 1, \ldots, n-1$). All bricks are identical with length (i.e., horizontal dimension) equal to 1. It is represented by a sequence of numbers, $(x_1, x_2, \ldots, x_n)$ where $x_i$ is the $x$-coordinate of the right edge of the $i$th brick. W.l.o.g., let $x_1 < x_2 < \cdots < x_n$. Recall from high school that the **center of gravity** (C.G.) of two masses $m$ and $M$, separated by distance $D$, is located at a point $p$ whose distance from $M$ is $mD/(m+M)$. The configuration $(x_1, \ldots, x_n)$ is **stable** if $n = 1$ or (recursively) $(x_2, \ldots, x_n)$ is stable, and the C.G. of $(x_2, \ldots, x_n)$ lies in the range $[x_1 - 1, x_1]$. Give a recursive formula for the position of the C.G. of $(x_1, \ldots, x_n)$.
$\diamondsuit$

**Exercise 6.9:** Let $c > 0$ be any real constant.
(a) Show that $\ln(n+c) - \ln n < c/n$ if $c < n$.
(b) Show that $|H_{x+c} - H_x| = \mathcal{O}(c/n)$ where $H_x$ is the generalized Harmonic function.
(c) Bound the sum $\sum_{i=1+\lfloor c \rfloor}^{n} \frac{1}{i(i-c)}$. $\diamondsuit$

**Exercise 6.10:** Consider the "alternating harmonic numbers",

$$A_n := \sum_{i=1}^{n} \frac{(-1)^{i+1}}{i} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots + \frac{(-1)^{n+1}}{n}. \tag{78}$$

(a) Show that $A_n$ converges to $\ln 2 = 0.693147\ldots$.
HINT: use the Maclaurin-Taylor series for $\ln(1-x)$.

(b) Use the fact that (see Chap.II,§19)

$$H_n = \ln n + \gamma + \frac{1}{2n} + o(1/n^2)$$

to conclude that $A_n = \ln 2 - (-1)^n \frac{1}{2n} + o(1/n^2)$.
HINT: Express an alternating harmonic number as the difference of two harmonic numbers $H_n$ and $H_{2n}$.

◇

**Exercise 6.11:** Consider $S_\infty(x)$ as a numerical sum.
(a) Prove that there is a unique value for $S_\infty(x)$ when $|x| < 1$.
(b) Prove that there are infinitely many possible values for $S_\infty(x)$ when $x \leqslant -1$.
(c) Are all real values possible as a solution to $S_\infty(-1)$? ◇

**Exercise 6.12:** Show the following useful estimate: $\ln(n) - (2/n) < \ln(n-1) < (\ln n) - (1/n)$. ◇

**Exercise 6.13:**
(a) Give the exact value of $\sum_{i=2}^{n} \frac{1}{i(i-1)}$. HINT: use partial fraction decomposition of $\frac{1}{i(i-1)}$.
(b) Conclude that $H_\infty^{(-2)} \leqslant 2$. ◇

**Exercise 6.14:** (Basel Problem) The goal is to give tight bounds for $H_n^{(-2)} := \sum_{i=1}^{n} \frac{1}{i^2}$ (cf. previous exercise).
(a) Let $S(n) = \sum_{i=2}^{n} \frac{1}{(i-1)(i+1)}$. Find the exact bound for $S(n)$.
(b) Let $G(n) = S(n) - H_n^{(-2)} + 1$. Now $\gamma' = G(\infty)$ is a real constant,

$$\gamma' = \frac{1}{1 \cdot 3 \cdot 4} + \frac{1}{2 \cdot 4 \cdot 9} + \frac{1}{3 \cdot 5 \cdot 16} + \cdots + \frac{1}{(i-1)\cdot(i+1)\cdot i^2} + \cdots .$$

Show that $G(n) = \gamma' - \theta(n^{-3})$.
(c) Give an approximate expression for $H_n^{(-2)}$ (involving $\gamma'$) that is accurate to $\mathcal{O}(n^{-3})$. Note that $\gamma'$ plays a role similar to Euler's constant $\gamma$ for harmonic numbers.
(d) What can you say about $\gamma'$, given that $H_\infty^{(-2)} = \pi^2/6$? Use a calculator (and a suitable approximation for $\pi$) to compute $\gamma'$ to 6 significant digits. ◇

**Exercise 6.15:** Show that $\sum_{i=1}^{r} \frac{x^i}{i} \leqslant x^{r+1} + \ln(r+1)$ where $r \in \mathbb{N}$ and $x > 0$. ◇

**Exercise 6.16:** Let $k \geqslant 1$ be a integer. We have the general formula $(1-x)^{-k} = \sum_{i \geqslant 0} x^i \binom{i+k-1}{k-1}$. Note that if $k = 1$, this is just the mother of series. Show this formula for $k = 2$ and $k = 3$. Generalize to all $k$. $\sum_{i\geqslant 0} x^i \binom{i+k-1}{k-1} = \sum_{i \geqslant 0} \frac{1}{(k-1)!} \frac{d^{k-1}}{dx^{k-1}} x^{i+k-1} = \ldots = \frac{1}{(1-x)^k}$. ◇

**Exercise 6.17:** Solve exactly (choose your own initial conditions):
(a) $T(n) = 1 + \frac{n+1}{n} T(n-1)$.
(b) $T(n) = 1 + \frac{n+2}{n} T(n-1)$. ◇

**Exercise 6.18:** Show that $\sum_{i=1}^n H_i = (n+1)H_n - n$. More generally,

$$\sum_{i=1}^n \binom{i}{m} H_i = \binom{n+1}{m+1}\left[H_{n+1} - \frac{1}{m+1}\right].$$

♢

**Exercise 6.19:** (J. van de Lune, 1980) Above, we defined $H_n := \sum_{i\geqslant 1}^n 1/i$ (descending sum). A variant that is neither a descending nor an ascending sum is to define $H(a,b) := \sum_{a\leqslant i\leqslant b} 1/i$ where the summation is over all integer values of $i$ in the range $[a,b]$. Show that this sum is satisfies

$$\sum_{a\leqslant x\leqslant b} \frac{1}{x} \leqslant \ln(b/a) + \min\{1, 1/a\}$$

♢

**Exercise 6.20:** Give a recurrence for $S_n^k$ (see (62)) in terms of $S_n^i$, for $i < k$. Solve exactly for $S_n^4$. ♢

**Exercise 6.21:** Derive the formula for the "geometric series of order 2", $k = 2$ in (70). ♢

**Exercise 6.22:** (a) Use Stirling's approximation to give an estimate of the exponent $E$ in the expression $2^E = \binom{2n}{n}$.
(b) (Feller) Show $\binom{2n}{n} = \sum_{k=0}^n \binom{n}{k}^2$. ♢

**Exercise 6.23:** Your architecture friend said that our brick tower design to achieve maximum overhang is unrealistic (we knew that). Here is a sequence of numbers that tend to infinity but slower: $G_n = \sum_{i=1}^n \frac{1}{i \lg i}$. Design a overhanging tower based on this sequence. What is the overhang formula $g(n)$ analogous to $h(n)$? ♢

**Exercise 6.24:** Manhattan has not been architecturally exciting since the Empire State Building (1931-present) and the Twin Towers of the World Trade Center (1973-2001). There is a new proposal to use the brick overhang design to build the most exciting building in the world. Each "brick" is a 100-foot cube. We want to have an overhang of 1000 feet. Question: *Use a pen-paper estimate for a lower bound on the height of this tower in miles.* One mile is 5280 feet, but you should approximate it as $2^{12} \approx 4000$ feet. The overhang formula of $h(n) = \frac{1}{2}H_{n-1}$ is found in a previous exercise.

♢

End Exercises

## §7. Standard Form and Summation Types

Recall that our goal is to reduce all recurrences to the **standard form**:

$$t(n) = t(n-1) + f(n). \tag{79}$$

We have noted that the solution is the descending sum

$$t(n) = S_f(n) = \sum_{i\geqslant 1}^n f(i) \tag{80}$$

assuming DIC with $t(n) = 0$ for $n < 1$. It is instructive to see this derived in a stylized way known as **telescopy**. Assuming the recurrence is valid for all $n \geqslant 1$, we have

$$\begin{array}{rcl} t(n) - t(n-1) & = & f(n) \\ t(n-1) - t(n-2) & = & f(n-1) \\ t(n-2) - t(n-3) & = & f(n-2) \\ & \vdots & \\ t(n-i) - t(n-i-1) & = & f(n-i) \\ & \vdots & \\ t(\{n\}+1) - t(\{n\}) & = & f(\{n\}+1) \end{array}$$

where $\{n\} = n - \lfloor n \rfloor$. Adding these $\lfloor n \rfloor$ equations together, we see that all but two terms on the left-hand side cancel ("telescoped"), leaving us

$$t(n) - t(\{n\}) = \sum_{i \geqslant 1}^{n} f(i).$$

By DIC, we may set $t(\{n\}) = 0$ to give us (80).

**¶28. The Euler-Maclaurin Approach to Summation.** What should we do if the open sum (80) does not reduce to one of the basic sums such as geometric or arithmetic series which we discussed in the previous section? Traditionally, the sum $S_f(n)$ is solved using the Euler-Maclaurin summation formula. This formula for ascending sums is as follows:

*view the sum $S_f(n)$ as a "discrete integral" of $f$*

$$\sum_{i=1}^{n-1} f(i) = \int_1^n f(x)dx + \left( \sum_{i=1}^{\infty} \frac{B_i f^{(i-1)}(x)}{i!} \right)_{x=1}^{x=n} \tag{81}$$

where $B_i$ is the $i$th Bernoulli number. See [7, p. 217]. The recursive formula for the Bernoulli numbers is given by $B_0 = 1$ and for $i \geqslant 1$,

$$B_i = -\sum_{k=0}^{i-1} \binom{i}{k} \frac{B_k}{i-k+1}.$$

The first few Bernoulli numbers are given by the following table:

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| $B_i$ | 1 | $-\frac{1}{2}$ | $\frac{1}{6}$ | 0 | $-\frac{1}{30}$ | 0 | $\frac{1}{42}$ | 0 | $-\frac{1}{30}$ |

In general, we have $B_i = 0$ if $i > 1$ is odd. How to use this formula for ascending sums? Assuming that we can integrate $f(x)$, and to take all its higher derivatives, we can just plug these values into the right hand side of (81). If $f(x)$ is a polynomial, then the right hand side is a finite sum. For instance, we obtain the Bernoulli-Faulhaber formulas for $S_n^k$ this way (Exercise). But in general, the right hand side is an infinite sum, and so we have to approximate it by truncation. Moreover, there are explicit integral formula for the remainder. As an upshot, we may obtain rather sharp bounds on (80), i.e., $S_f(n) = F(n)\left[1 + o(n)\right] \sim F(n)$.

The assumption that we can integrate $f(x)$ is a nontrivial assumption; you may recall some integral formulas from elementary calculus. In computer algebra, Robert Risch (1968) gave an algorithm to decide when an elementary[16] function $f(x)$ has another elementary function as integral. The algorithm is complete modulo the ability to decide if an elementary constant is equal to zero. Unfortunately, the latter is a deep open problem.

[16] "Elementary functions" are sometimes informal. But we may also define it as follows: they are the constants $0, 1$ or identity $f(x) = x$, and are closed under the four arithmetic operations, and under taking exponentials, logarithms, radicals, and trigonometric functions. An elementary constant is just an elementary function constructed from 0 and 1 alone.

**¶29. Polynomial-type and Exponential-type Sums.** Instead of the sharp bounds afforded by the Euler-Maclaurin formula, suppose we want to determine a descending sum $S_f(n) = \sum_{i\geqslant 1}^{n} f(i)$ up to $\Theta$-order:

$$S_f(n) = \Theta(F(n)) \tag{82}$$

for some explicit function $F(n)$. We show that such weaker bounds are easily obtained for a large class of functions $f(n)$. We use only elementary arguments, avoiding calculus. We first make some initial observation based on the form of the solution (82). The form already says that $S_f(n)$ must be eventually non-negative. So it is natural to require

*No calculus please, we are computer scientists!*

$$f > 0 \text{ (ev.) and } f(n) = \downarrow \text{ (ev.)}. \tag{83}$$

Second, we require a suitable finiteness property: it is best to motivate this with an example. If $f(x) = 1/(x-2)$, and $0 < \varepsilon < 1$, then

$$S_f(3+\varepsilon) = \frac{1}{1+\varepsilon} + \frac{1}{\varepsilon} - \frac{1}{1-\varepsilon}$$

Since $f(2)$ is undefined, our convention says that when $f(2)$ appears in the sum $S_f(x)$, we take its value to be 0. Thus $S_f(3) = 1 + 0 - 1 = 0$, and $S_f(3+\varepsilon)$ is arbitrarily large (as $\varepsilon \to 0$) or arbitrarily small (as $\varepsilon \to 1$). To avoid such behavior, we define the following property: $f$ is **locally bounded** if

$$(\forall x_0 > 0)(\exists C > 0)(\forall x)[0 < x < x_0 \Rightarrow \ |f(x)| \leqslant C\ ]. \tag{84}$$

For example, the function $1/(x-2)$ is not locally bounded. If $f$ is locally bounded, then for all $x > 0$, there is a neighborhood $N$ of $x$ and $C > 0$ such that $|f(y)| \leqslant C$ for $y \in N$. Note that a locally bounded function $f$ can be unbounded as $x \to \infty$, or if $x \leqslant 0$.

We next introduce two "growth types" in complexity functions $f$:

**Polynomial Type:** A real function $f$ is **polynomial-type** if $f \geqslant 0$ (ev.), is non-decreasing (ev.) and there is some $C > 1$ such that

$$f(x) \leqslant C \cdot f(x/2) \text{ (ev.)}.$$

For example, the function $f(x) = x^2$ is polynomial-type because $x^2 \leqslant C \cdot (x/2)^2$ if we choose $C \geqslant 4$. Note that $f(x) \leqslant Cf(x/2) \leqslant C^2 f(x/4) \leqslant \cdots \leqslant C^k f(x/2^k)$. Choosing $k = \lceil \lg x \rceil$,

$$f(x) = O(C^k) = O(x^{\lg C}).$$

Hence, each polynomial-type function is bounded by a polynomial. Here are more examples of polynomial-type functions (assuming $a \geqslant 0$):

$$f_0(x) = x^a, \quad f_1(x) = \log x, \quad f_2(x) = f_0(x) f_1(x), \quad f_3(x) = (f_0(x))^a. \tag{85}$$

We also note a negative example: $f(x) = 1/x$ is not polynomial-type because it is decreasing.

**Exponential Type:** The function $f$ is **exponential-type** if it increases exponentially or it decreases exponentially:

(a) $f$ **increases exponentially** if there exists real numbers $C > 1$ and $k > 0$ such that

$$f(x) \geqslant C \cdot f(x-k) \text{ (ev.)}. \tag{86}$$

For example, the function $f(x) = 2^x$ increases exponentially because $2^x \geqslant C2^{x-1}$ if we choose $k = 1$ and $C = 2$. Again, $f(x) = 2^{2^x}$ increases exponentially because $2^{2^x} = \left(2^{2^{x-1}}\right)^2 \geqslant C \cdot 2^{2^{x-1}}$ if we choose $C = 2$ and $k = 1$ (for $x \geqslant 1$). Here are more examples (assuming $b > 1$):

$$g_0(x) = b^x, \quad g_1(x) = x!, \quad g_2(x) = g_0(x) g_1(x), \quad g_3(x) = b^{g_0(x)} \tag{87}$$

(b) $f$ **decreases exponentially** if there exists real numbers $C > 1$ and $k > 0$ such that

$$f(x) < f(x-k)/C \text{ (ev.)}. \tag{88}$$

For example, the function $f(x) = 2^{-x}$ decreases exponentially because $2^{-x} \leqslant 2^{-(x-1)}/C$ if we choose $k = 1$ and $C = 2$. The following are further examples (assuming $b > 1$):

$$h_0(x) = b^{-x}, \quad h_1(x) = x^{-x}, \quad h_2(x) = h_0(x) h_1(x), \quad h_3(x) = b^{h_0(x)} \tag{89}$$

It follows immediately from the definitions (86) and (88) that $f$ *is increasing exponentially iff* $1/f$ *is decreasing exponentially.* In proofs, we can usually take $k = 1$ in the definition of exponential-types: i.e., if $g(n)$ is increasing exponentially, $g(n) \geqslant Cg(n-1)$ and if $h(n)$ is decreasing exponentially, $h(n) < h(n-1)/C$.

We say that the descending sum $S_f(n) := \sum_{x\geqslant 1}^n f(x)$ is **polynomial-type** or **exponential-type**, according to the type of $f$. The next theorem gives a simple rule for bounding such sums.

**Theorem 6 (Summation Rules)** *Assume $f$ is eventually positive and well-defined (83), and is locally bounded (84). Then*

$$S_f(n) = \Theta \begin{cases} nf(n) & \text{if } f \text{ is polynomial-type,} \\ f(n) & \text{if } f \text{ is increasing exponential-type,} \\ 1 & \text{if } f \text{ is decreasing exponential-type.} \end{cases}$$

*Proof.* All the following arguments hold eventually.
CASE (i): For a polynomial-type sum, using the fact that $f$ is eventually non-decreasing, we get the upper bound $S_f(n) \leqslant \sum_{x\geqslant 1}^n f(n) = \mathcal{O}(nf(n))$. For lower bound, we also need that $f(x) \leqslant Cf(x/2)$ (ev.) for some $C > 0$:

$$\begin{aligned} S_f(n) &\geqslant \sum_{x\geqslant n/2}^n f(x) \\ &\geqslant \sum_{x\geqslant n/2}^n f(n/2) \geqslant \lceil n/2 \rceil f(n/2) \\ &\geqslant \lceil n/2 \rceil \frac{f(n)}{C} = \Omega(nf(n)). \end{aligned}$$

CASE (ii-a): For an increasing exponential sum, there is some $C > 1$, $k > 0$ and $n_0 > 0$ such that for all $n \geqslant n_0$, we have $f(n) \geqslant Cf(n-k)$. For any $n > n_0$, let $j = \lceil (n-n_0)/k \rceil$ and thus $n_0 \geqslant n - jk > n_0 - k$. Also $S_f(n-jk) \leqslant K$ (for some constant $K > 0$) since $f$ is locally bounded. Thus

$$\begin{aligned} S_f(n) &= [f(n) + f(n-k) + f(n-2k) + \cdots + f(n-(j-1)k)] + S_f(n-jk) \\ &\leqslant f(n)\left[1 + \tfrac{1}{C} + \tfrac{1}{C^2} + \cdots\right] + S_f(n-jk) \\ &= f(n)\tfrac{C}{C-1} + K \\ &= \mathcal{O}(f(n)) \qquad \text{(since } f > 0 \text{ (ev.)).} \end{aligned}$$

Since $S_f(n) = \Omega(f(n))$, we conclude that $S_f(n) = \Theta(f(n))$.
CASE (ii-b): For a decreasing exponential sum, there is some $C > 1$, $k > 0$ and $n_0 > 0$ such that for all $n \geqslant n_0$, we have $Cf(n) \leqslant f(n-k)$. As before, let $n > n_0$ and $j = \lceil (n-n_0)/k \rceil$. Then $n_0 \geqslant n - jk > n_0 - k$. We may assume that for all $x \in (n_0 - k, n_0]$, $f(x) \geqslant \varepsilon$ (for some $\varepsilon > 0$). Moreover, $S_f(n-(j+1)k) \leqslant K$ for some constant $K$ because $f$ is locally bounded.

$$\begin{aligned} S_f(n) &= S_f(n-(j+1)k) + [f(n-jk) + f(n-(j-1)k) + \cdots + f(n-k) + f(n)] \\ &\leqslant K + f(n-jk)\left[1 + \tfrac{1}{C} + \tfrac{1}{C^2} + \cdots\right] \\ &\preceq f(n-jk) \qquad \text{(since } f(n-jk) \geqslant \epsilon\text{).} \end{aligned}$$

Since $S_f(n) \geqslant f(n-jk) \geqslant \epsilon$, we conclude that $S_f(n) = \Theta(1)$. **Q.E.D.**

Let us apply this theorem to determine the $\Theta$-order of various sum. Once we know the type of the sum, it is a simple matter of writing down the solution:

- Polynomial Sums (recall (85))

$$\sum_{i\geqslant 1}^n i \log i = \Theta(n^2 \log n), \qquad \sum_{i\geqslant 1}^n \log i = \Theta(n \log n), \qquad \sum_{i\geqslant 1}^n i^a = \Theta(n^{a+1}) (a \geqslant 0). \qquad (90)$$

- Exponentially Increasing Sums (recall (87))

$$\sum_{i\geqslant 1}^{n} b^i = \Theta(b^n)(a \geqslant 0), \qquad \sum_{i\geqslant 1}^{n} i^{-5}2^{2^i} = \Theta(n^{-5}2^{2^n}), \qquad \sum_{i\geqslant 1}^{n} i! = \Theta(n!) \quad . \tag{91}$$

- Exponentially Decreasing Sums (recall (89))

$$\sum_{i\geqslant 1}^{n} b^{-i} = \Theta(1)(a \geqslant 0), \qquad \sum_{i\geqslant 1}^{n} i^2 i^{-i} = \Theta(1), \qquad \sum_{i\geqslant 1}^{n} i^{-i} = \Theta(1) \quad . \tag{92}$$

**¶30. Reducing to summations we can bound.** A summation that does not fit the framework of Theorem 6 can sometimes be reduced to one that does. A simple case is when summation does not begin from $i = 1$. As another example, consider

$$S := \sum_{i\geqslant 1}^{n} \frac{i!}{\lg^i n}, \tag{93}$$

which has terms depending on $i$ as well as on the limit $n$. Write $S = \sum_{i\geqslant 1}^{n} f(i,n)$ where

$$f(i,n) = \frac{i!}{\lg^i n}.$$

We note that $f(i,n)$ is increasing exponentially for $i \geqslant 2\lg n$ (ev. $n$), since $f(i,n) = \frac{i}{\lg n} f(i-1,n) \geqslant 2f(i-1,n)$. Hence we may split the summation into two parts, $S = A + B$ where $A$ comprise the terms for which $i < 2\lg n$ and $B$ comprising the rest. Since $B$ is an exponential sum, we have $B = \Theta(f(n,n))$. We can easily use Stirling's estimate for $A$ to see that $A = \mathcal{O}(\log^{3/2} n) = \mathcal{O}(f(n,n))$. Thus $S = \Theta(f(n,n))$.

**¶31. A Counter Example.** A function cannot be both polynomial-type and exponential type (Exercise). Many common functions we encounter will be either polynomial-type or exponential-type. We now show a function that is neither:

**Lemma 7** *The function $f(n) = n^{\ln n}$ is neither polynomial-type nor exponential-type.*

*Proof.* Showing that $f(n)$ is not polynomial-type is easy: the ratio

$$f(n)/f(n/2) = n^{\lg(n)}/n^{(\lg(n/2))} \cdot 2^{\lg n/2} = n^2/2$$

is unbounded, so $f$ is not polynomial-type.

To show that it is not exponential-type, assume by way of contradiction that there exists $C_0 > 1$ such that

$$f(n) \geqslant C_0 f(n-1) \text{ (ev.)}. \tag{94}$$

We use a well-known bound (see Appendix) says that for $|x| < 1$,

$$\ln(1+x) < x. \tag{95}$$

Also from (72) and (73), we conclude that

$$\ln n + \gamma \leqslant H_n \leqslant \ln n + \gamma + (1/n) \quad \text{(ev.)}. \tag{96}$$

The following inequalities hold eventually:

$$\begin{aligned} \ln n &\leqslant H_n - \gamma \\ &\leqslant (1/n) + \ln(n-1) + (1/n) \\ &= \ln(n-1) + (2/n). \end{aligned} \tag{97}$$

We now get a contradiction:

$$\begin{array}{rcll} f(n) &=& \left[(n-1)(1+\frac{1}{n-1})\right]^{\ln n} & \\ &\leqslant& (n-1)^{\ln(n-1)+(2/n)}(1+\frac{1}{n-1})^{\ln n} & \text{(by (97))} \\ &=& f(n-1)\cdot(n-1)^{2/n}\cdot 2^{\ln(1+\frac{1}{n-1})\ln n} & \\ &\leqslant& f(n-1)\cdot 2^{2\ln(n-1)/n}\cdot 2^{\frac{\ln n}{n-1}} & \text{(by (95))} \\ &=& f(n-1)\cdot C_1(n) & \end{array}$$

where $C_1(n) := 2^{2\ln(n-1)/n}\cdot 2^{\frac{\ln n}{n-1}}$. Since $\ln C_1(n) = (2\ln(n-1)/n) + (\ln n/(n-1)) \to 0$ as $n \to \infty$, we conclude that $C_1(n) \leqslant C_0$ (ev.). This show $f(n) \leqslant f(n-1)C_0$ (ev.), contradicting (94). **Q.E.D.**

How do we estimate the sum $S_f(n) := \sum_{x\geqslant 0}^{n} f(x)$ since we cannot apply Theorem 6 when $f$ is neither polynomial- nor exponential-type? In this case, techniques similar to polynomial and exponential sums still give reasonably tight bounds (but not $\Theta$-order): $f(n) \leqslant S_f(n) \leqslant nf(n) \leqslant f(n)^{1+\varepsilon}$ for any $\varepsilon > 0$.

**¶32. Closure Properties: How to recognize growth types** To apply the summation rules of Theorem 6, we want to rapidly classify functions according to their growth types. For this purpose, we can use our next lemma which shows that these growth types are closed under various operations.

**Lemma 8 (Closure Properties)** *Let $a \in \mathbb{R}$.*

*(a) Polynomial-type functions are closed under addition, multiplication, and raising to any positive power $a > 0$.*

*(b) Exponential-type functions $f$ preserve their subtypes under addition, multiplication and raising to any power $a$. In case $a > 0$, the function $f^a$ will not change its subtype (increasing or decreasing). In case $a < 0$, the function $f^a$ will change its subtype.*

*(c) If $f$ is polynomial-type and $\lg f$ is non-decreasing then $\lg f$ is also polynomial-type. If $f$ is exponential-type and $a > 1$ then so is $a^f$.*

*(d) If $f$ is exponentially increasing-type, then its type is preserved by multiplication with a non-decreasing function (e.g., polynomial-type).*

*Proof.* All the inequalities in the following proofs are assumed to hold eventually:

(a) Assume $f(n) \leqslant Cf(n/2)$ and $g(n) \leqslant Cg(n/2)$ for some $C > 1$. Then $f(n)+g(n) \leqslant C(f(n/2)+g(n/2))$, $f(n)g(n) \leqslant C^2 f(n/2)g(n/2)$, and for any $e > 0$, $f(n)^e \leqslant C^e f(n/2)^e$.

(b) Assume $g_i(n) \geqslant Cg_i(n-1)$ and $h_i(n) \leqslant ch_i(n-1)$. for some $C > 1, c < 1$, and for $i = 0, 1$. Also, let $g = g_0$, $h = h_0$. Closure under addition: $g_0(n)+g_1(n) \geqslant C(g_0(n-1)+g_1(n-1))$ and $h_0(n)+h_1(n) \leqslant c(h_0(n-1)+h_1(n-1))$. Closure under product: $g_0(n)g_1(n) \geqslant C^2 g_0(n-1)g_1(n-1))$ and $h_0(n)h_1(n) \leqslant c^2 h_0(n-1)h_1(n-1)$. Closure under raising to power $e$: If $e > 0$, then $g^e(n) \geqslant C^e g^e(n-1)$ and $h^e(n) \leqslant c^e h^e(n-1)$ where $C^e > 1$ and $c^e < 1$. If $e < 0$, then $g^e(n) \leqslant C^e g^e(n-1)$ and $h^e(n) \geqslant c^e h^e(n-1)$ where $C^e < 1$ and $c^e > 1$.

(c) If $f$ is polynomial-type, then $\log(f(n)) \leqslant (\log C) + \log(f(n/2)) \leqslant (1 + (\log C)/c)\log(f(n/2))$, where $\log(f(n/2)) \geqslant c > 0$ for some constant $c$. This proves $\log f$ to be polynomial-type. If $g, h$ is exponential type as in (b), then note that $Cg(n) \geqslant (C-1) + g(n)$ since $g(n) \geqslant 1$. Thus

$$\begin{array}{rcl} b^{g(n)} &\geqslant& b^{Cg(n-1)} \geqslant b^{(C-1)+f(n-1)} \\ &\geqslant& b^{C-1}2^{f(n-1)}. \end{array}$$

(d) Let $f(n) \geqslant Cf(n-1)$, $C > 1$ and $g$ be a non-decreasing function. Then $f(n)g(n) \geqslant Cf(n-1)g(n-1)$.

**Q.E.D.**

**¶* 33. Generalized Harmonic Numbers and Trichotomy.** For all $n, \alpha \in \mathbb{R}$, define[17] the **generalized harmonic number**

$$\begin{aligned} H^{(\alpha)}(n) &:= \sum_{x \geqslant 1}^{n} x^{\alpha} \\ &= n^{\alpha} + (n-1)^{\alpha} + (n-2)^{\alpha} + \cdots + (\{n\}+1)^{\alpha}, \end{aligned} \tag{98}$$

written as a descending sum (57). Note that $H^{(\alpha)}(n) = 0$ for $n < 1$. The harmonic numbers $H_n$ is just the special case of $\alpha = -1$ and $n$ is a positive integer. The arithmetic series in ¶23 corresponds to $H^{(\alpha)}(n)$ where $\alpha \in \mathbb{N}$. When $\alpha \leqslant -1$, the sum $H^{(\alpha)}(n)$ is bounded as $n \to \infty$; the limiting value $H^{(\alpha)}(\infty)$ is the value of the Riemann zeta function at $-\alpha$: $\zeta(\alpha) := \sum_{i=1}^{\infty} n^{-\alpha} = H^{(-\alpha)}(\infty)$. For instance, $\zeta(2) = H^{(-2)}(\infty) = \pi^2/6$. An Exercise estimates the sum $H^{(-2)}(n)$. Just as Euler's constant $\gamma$ arise in estimates of $H^{(-1)}(n)$, an analogous constant arise in estimating $H^{(-2)}(n)$. The following lemma determines the $\Theta$-order of $H^{(\alpha)}(n)$ for fixed $\alpha$:

$\zeta(3) = H^{(-3)}(\infty) = 1.202056903159594$ *is Apéry's constant, provably irrational.*

**Theorem 9 (Trichotomy)** *For all* $\alpha \in \mathbb{R}$*,*

$$H^{(\alpha)}(n) = \Theta \begin{cases} 1 & \text{if } \alpha < -1 \\ \lg n & \text{if } \alpha = -1 \\ n^{\alpha+1} & \text{if } \alpha > -1 \end{cases} \tag{99}$$

*“proof” that* $\ln x$ *is identically* $1$*:* $\frac{d(\ln x)}{dx} = \frac{1}{x}$ and $\frac{d(1)}{dx} = \frac{d(x^0)}{dx} = x^{0-1} = \frac{1}{x}$. So $\ln x = 1$.

*Proof.* It is best to initially assume $n+1$ is a power of 2. Then

$$\begin{aligned} H^{(\alpha)}(n) &= \sum_{k=1}^{\lg(n+1)} \left( \sum_{i=2^{k-1}}^{2^k - 1} i^{\alpha} \right) \\ &\overset{(*)}{=} \sum_{k=1}^{\lg(n+1)} 2^k \cdot \Theta\left(2^{k\alpha}\right) \\ &= \sum_{k=1}^{\lg(n+1)} \Theta\left(2^{k(1+\alpha)}\right). \end{aligned} \tag{100}$$

The first summation is a direct analogy with (74), the technique of splitting a sum into about $\lg n$ groups. Note that the slick use of $\Theta$ in step $(*)$ to capture upper and lower bounds simultaneously. Let us spell this out: note that $i = C \cdot 2^k$ for some $C \in [\frac{1}{2}, 1]$. Therefore $i^{\alpha} = C' \cdot 2^{k\alpha}$ for some $C' \in [2^{-\alpha}, 1]$ (if $\alpha \geqslant 0$) or $C' \in [1, 2^{-\alpha}]$ (if $\alpha < 0$). Finally, we give a closed form for the open sum (100): notice that if $1 + \alpha = 0$ then we trivially get

$$\sum_{k=1}^{\lg(n+1)} 2^{k(1+\alpha)} = \Theta(\lg(n+1)).$$

If $1 + \alpha < 0$, then the sum is decreasing exponentially and Theorem 6 yields

$$\sum_{k=1}^{\lg(n+1)} 2^{k(1+\alpha)} = \Theta(1).$$

[17] Knuth [11, p. 74] writes $H_n^{(-\alpha)}$ for our $H^{(\alpha)}(n)$. Presumably, his definition assumes that $n$ is integer.

If $1+\alpha > 0$, then the sum is increasing exponentially and Theorem 6 yields

$$\sum_{k=1}^{\lg(n+1)} 2^{k(1+\alpha)} = \Theta\left((2^{\lg(n+1)})^{1+\alpha}\right) = \Theta\left(n^{1+\alpha}\right).$$

When $n+1$ is not a power of 2, we can replace $n$ by $\overline{n} = 2^{\lceil \lg(n+1)\rceil} - 1$ and $\underline{n} = 2^{\lfloor \lg(n+1)\rfloor} - 1$ for upper and lower bounds, giving the same $\Theta$-bound. **Q.E.D.**

This result has two significance. First, up to $\Theta$-order, the summation (99) unifies the standard bounds for the arithmetic series (62), harmonic numbers (72). Indeed, the proof also shows an intimate connection to geometric sums (64). That is, after the grouping transformation $(*)$, $H^{(\alpha)}(n)$ becomes a geometric sum. Second, the solution of $H^{(\alpha)}(n)$ is based on a trichotomy that will be repeated in the Master Theorem below. Although formula (99) has an analogue in calculus, our proof uses only elementary arguments. In [19], we generalize the transformation $(*)$ to any descending sum $S_f$ where $f$ belongs to the class of "exponential-logarithmic functions" (or EL-functions [19]). The result says that if $f$ is an EL-function, then $S_f$ is $\Theta$-order of another EL-function.

*The trichotomy motif!*

Application: to solve the recurrence $T(n) = 2T(n/2) + (n/\lg n)$, we convert it to the standard form

$$t(N) := t(N-1) + 1/N \tag{101}$$

using the substitution $t(N) = T(2^N)/2^N$, where $N = \lg n$ is a real variable. According to (60), $t(N) = H^{(-1)}(N)$. Back solving, the original recurrence has solution $T(n) = nH^{(-1)}(\lg n) = \Theta(n \ln \lg n)$.

**¶34. Grouping: Breaking Up into Big and Small Parts.** The above example (93) illustrates the technique of breaking up a sum into two parts, one containing the "small terms" and the other containing the "big terms". This is motivated by the wish to apply different summation techniques for the 2 parts, and this in turn determines the cutoff point between small and big terms. Suppose we want to show

$$H_n = \sum_{i\geqslant 1}^{n} \frac{1}{i} = O(\sqrt{n}).$$

Break $H_n$ into two summations, $H_n = A_n + B_n$ where

$$A_n = \sum_{i\geqslant 1}^{n-\lfloor n-\sqrt{n}\rfloor} \frac{1}{i}$$

comprises the "big terms" (there are at most $\sqrt{n}$ terms in $A_n$), and $B_n$ contains the remaining $\lfloor n-\sqrt{n}\rfloor$ "small terms". Then

$$A_n \leqslant \sum_{i\geqslant 1}^{n-\lfloor n-\sqrt{n}\rfloor} \frac{1}{i} \leqslant \sqrt{n}$$

and

$$B_n = \sum_{i\geqslant n-\lfloor n-\sqrt{n}\rfloor}^{n} \frac{1}{i} \leqslant \sum_{i=1}^{n} \frac{1}{\sqrt{n}} = \sqrt{n}.$$

Thus $S_n \leqslant 2\sqrt{n} = O(\sqrt{n})$ as desired.

We can generalize the grouping idea to prove the following:

$$H_n < kn^{1/k} \tag{102}$$

for any integer $k \geqslant 2$. We break the summation $H_n$ into $k$ subsums, $H_n = A_n(1) + A_n(2) + \cdots + A_n(k)$ where $A_n(1)$ comprises the first $\lceil n^{1/k} \rceil$ terms of $H_n$, $A_n(2)$ comprises the next $\lceil n^{2/k} \rceil - \lceil n^{1/k} \rceil$ terms, etc, where in general, $A_n(j)$ comprises the next $\lceil n^{j/k} \rceil - \lceil n^{(j-1)/k} \rceil$ terms. It is easy to see that each $A_n(j)$ is bounded by $n^{1/k}$ and this proves (102). This proves that $H_n$ is $O(n^c)$ for any $c > 0$. This also implies

$$H_n = o(n^c), \qquad \log_b n = o(n^c).$$

Exercises

**Exercise 7.1:** Show that no function can be both polynomial-type and exponential type. ♢

**Exercise 7.2:** Show that $n/\ln n$ is non-decreasing. ♢

**Exercise 7.3:** (a) Let $f(x) = 4x^2 - 10x$. Prove that $f(x)$ is polynomial-type.
Note that you must show two implicit constants, $C > 1$ and $x_0 > 0$. For full credit, *choose the smallest integer value of $C$, and subject to this, the smallest value of $x_0$.*

(b) Let $f(x) = 2^x - 10x$. Prove that $f(x)$ is exponentially increasing-type.
By definition, you must show 3 constants $C > 1$, $k > 0$ and $x_0 > 0$ such that

$$(\forall x > x_0)[f(x) > C \cdot f(x - k)] \tag{103}$$

Please choose $k = 1$, then choose $C$ to be the smallest possible integer, and then choose $x_0$ to be the smallest possible integer.

(c) Let $f(x) = 2^x/(x!)$. Prove that $f(x)$ is exponentially decreasing-type.
Similar to part(b), we need 3 constants $C > 1$, $k > 0$ and $x_0 > 0$. Again, please set $k = 1$ and choose smallest possible integer values of $C$ and $x_0$.

(d) Prove that polynomial-types are closed under multiplication: i.e., if $f(x) \geqslant 0$ and $g(x) \geqslant 0$ are polynomial-types, so is $h(x) := f(x)g(x)$.

(e) Please state the $\Theta$-order of the following sums:

(i) $\sum_{i\geqslant 1}^{n} i^7 \log^3 i$

(ii) $\sum_{i\geqslant 1}^{n} i^2 \log^{-1} i$

(iii) $\sum_{i\geqslant 1}^{n} \frac{2^i}{i^5 \log i}$

Justify your derivations – you may quote any facts in these lecture notes (please be specific!). Otherwise you must prove it yourself!

♢

**Exercise 7.4:** For each function, determine its growth type. You may use closure properties (Lemma 8) in the text; if necessary, argue from first principles (calculus is not necessary but may simplify your argument):
(a) $2^{n^2}$, (b) $(\lg\lg n)^2$, (c) $n/\log n$. ♢

**Exercise 7.5:** Verify that the examples in (90), (91) and (92) are, indeed, as claimed, polynomial type or exponential type. ♢

**Exercise 7.6:** Let $T_n$ be a perfect binary tree with $n \geqslant 1$ nodes. So $n = 2^{h+1} - 1$ where $h$ is the height of $T_n$. Suppose an algorithm has to visit all the nodes of $T_n$ and at each node of height $i \geqslant 0$, expend $(i+1)^2$ units of work. Let $T(n)$ denote the total work expended by the algorithm at all the nodes. Give a tight upper and lower bounds on $T(n)$. ♢

**Exercise 7.7:** (a) Show that the summation $\sum_{i\geqslant 2}^{n} (\lg n)^{\lg n}$ is neither polynomial-type nor exponential-type.
(b) Estimate this sum. ♢

**Exercise 7.8:** Give the $\Theta$-order bound of these sums using our summation rules for each growth type. Be sure to justify why a function has a specific growth type.
(a) $S(n) = \sum_{i=1}^{n} i^2(6i^2 - 3i + 2)(i + 4)$
(b) $S(n) = \sum_{i=1}^{n} 2^i (3i + 1)^2 \log^3 i$
(c) $S(n) = \sum_{i=1}^{n} \frac{2^i}{i!} i^2$ ♢

**Exercise 7.9:** For this problem, please use elementary estimates (arguments from first principles). Show that $H_n = o(n^\alpha)$ for any $\alpha > 0$. HINT: Generalize the argument in the text. ♢

**Exercise 7.10:** Use the method of grouping to show that $S(n) = \sum_{i=1}^{n} \frac{\lg i}{i}$ is $\Omega(\lg^2 n)$. ♢

**Exercise 7.11:** Give the $\Theta$-order of the following sums. If you use our summation rules, then you must show that the terms has the appropriate growth types.
(a) [cf. Knuth, vol.1, p.43] $S = \sum_{i=1}^{n} \sqrt{i}$.
(b) $S = \sum_{i=1}^{n} \lg(n/i)$. ♢

**Exercise 7.12:** Let $f(i) = f_n(i) = \frac{i-1}{n-i+1}$. The sum $F(n) = \sum_{i=1}^{n} f_n(i)$ is neither polynomial-type nor exponential-type. Give a $\Theta$-order bound on $F(n)$. HINT: transform this into something familiar.
♢

**Exercise 7.13:** Can our summation rules for $S(n) = \sum_{i=1}^{n} f(i)$ be extended to the case where $f(i)$ is "decreasing polynomially", suitably defined? NOTE: such a definition must somehow distinguish between $f(i) = 1/i$ and $f(i) = 1/(i^2)$, since in one case $S(n)$ diverges and in the other it converges as $n \to \infty$. ♢

End Exercises

## §8. Domain Transformation

So our goal for a general recurrence is to transform it into the standard form. You may think of change of domain as a "change of scale". Transforming the domain of a recurrence equation may sometimes bring it into standard form. Consider

$$T(N) = T(N/2) + N. \tag{104}$$

We define

$$t(n) := T(2^n), \quad N := 2^n.$$

This transforms the original $N$-domain into the $n$-domain. The new recurrence is now in standard form,

$$t(n) = t(n-1) + 2^n.$$

By DIC, we may choose the boundary condition $t(n) = 0$ for all $n < 0$, we get the descending sum $t(n) = \sum_{i\geqslant 0}^{n} 2^i$. Writing $b = n - \lfloor n \rfloor = \{n\}$, we transform it into an ascending sum $t(n) = \sum_{j=0}^{n-b} 2^{n-j} = 2^b \sum_{j=0}^{n-b} 2^j$ (why?). We know how to sum $\sum_{j=0}^{n-b} 2^j$ as $2^{n+1-b} - 1$ and thus $t(n) = 2^b(2^{n+1-b} - 1) = 2^{n+1} - 2^b$; hence, $T(N) = 2N - 2^b$.

Note the payoff in our decision of "going real" in solution of recurrences: the transformed function $t(n)$ is real if $T(N)$ is real. But if $T(N)$ were integer, $t(n)$ would not remain integer.

**¶35. Logarithmic transform.** More generally, consider the recurrence

$$T(N) = T\left(\frac{N}{c} - d\right) + F(N), \quad c > 1, \tag{105}$$

and $d$ is an arbitrary constant. It is instructive to begin with the case $d = 0$. Consider the "logarithmic transformation" of the argument $N$ to the new argument $n := \log_c(N)$. Then $N/c$ transforms to $\log_c(N/c) = n - 1$. Then $T(N) = T(N/c) + F(N)$ transforms into the new recurrence

*So $N = c^n$*

$$t(n) = t(n-1) + f(n)$$

where we define

$$t(n) := T(c^n) = T(N), \qquad f(n) := F(N).$$

The preceding manipulation exploits some implicit conventions: $N \leftrightarrow n$, $T \leftrightarrow t$, $F \leftrightarrow f$. This might be confusing in more complicated situations, so let us make the connection between $t$ and $T$ more explicit. Let $\tau$ denote the **domain transformation function**,

$$\tau(N) := \log_c(N), \qquad \tau^{-1}(n) = c^n$$

Then $t(\tau(N))$ is defined to be $T(N)$, valid for large enough $N$. In order for this to be well-defined, we need $\tau$ to have an inverse for large enough $n$. Then we can write

*"$n$" is a short-hand for "$\tau(N)$"*

$$t(n) := T(\tau^{-1}(n)).$$

We now return to the general case where $d$ is an arbitrary constant. Note that if $d < 0$ then we must assume that $N$ is sufficiently large (how large?) so that the recurrence (105) is meaningful (*i.e.*, $(N/c) - d < N$). The following "generalized logarithmic transformation"

$$n := \tau(N) = \log_c(N + \frac{cd}{c-1}) \tag{106}$$

will reduce the recurrence to standard form. To see this, note that the inverse transformation is

$$\begin{aligned} N &:= c^n - \frac{cd}{c-1} \\ &= \tau^{-1}(n) \\ (N/c) - d &= c^{n-1} - \frac{d}{c-1} - d \\ &= c^{n-1} - \frac{cd}{c-1} \\ &= \tau^{-1}(n-1). \end{aligned}$$

Writing $t(n)$ for $T(\tau^{-1}(n))$ and $f(n)$ for $F(\tau^{-1}(n))$, we convert equation (105) to

$$\begin{aligned} t(n) &= T(\tau^{-1}(n)) && \text{(by definition of } t(n)\text{)} \\ &= T(N) && (N = \tau^{-1}(n)) \\ &= T((N/c) - d) + F(N) && \text{(expansion)} \\ &= T(\tau^{-1}(n-1)) + F(\tau^{-1}(n)) && \text{(domain transform)} \\ &= t(n-1) + f(n) && \text{(definition of } t(n) \text{ and } f(n)\text{)} \\ &= \textstyle\sum_{i\geqslant 1}^{n} f(i) && \text{(telescopy and by DIC)} \end{aligned}$$

To finally "solve" for $t(n)$ we need to know more about the function $F(N)$. For example, if $F(N)$ is a polynomially bounded function, then $f(n) = F(c^n - \frac{cd}{c-1})$ would be $\Theta(F(c^n))$. This is the justification for ignoring the additive term "$d$" in the equation (105).

**¶36. Division transform.** Notice that the logarithmic transform case does not quite capture the following closely related recurrence

$$T(N) = T(N-d) + F(N), d > 0. \tag{107}$$

It is easy to concoct the necessary domain transformation: replace $N$ by $n = N/d$ and substituting

$$t(n) = T(dn)$$

will transform it to the standard form,

$$t(n) = t(n-1) + F(dn).$$

Again, we can explicitly introduce the "division transform" function $\tau(N) = N/d$, etc.

**¶37. General Pattern.** In general, we consider $T(N) = T(r(N)) + F(N)$ where $r(N) < N$ is some function. We want a domain transform $n = \tau(N)$ so that

$$\tau(r(N)) = \tau(N) - 1. \tag{108}$$

The generalized logarithm transform (106) is of this type. Here is another example: if $r(N) = \sqrt{N}$ we may choose

$$\tau(N) = \lg\lg(N). \tag{109}$$

Then we see that

$$\tau(\sqrt{N}) = \lg(\lg(\sqrt{N})) = \lg(\lg(N)/2) = \lg\lg N - 1 = \tau(N) - 1.$$

Applying this transformation to the recurrence

$$T(N) = T(\sqrt{N}) + N, \tag{110}$$

we may define $t(n) := T(\tau^{-1}(n)) = T(2^{2^n}) = T(N)$, thereby transforming the recurrence (110) to to $t(n) = t(n-1) + 2^{2^n}$.

Note that the transformation (109) may be regarded as two applications of the logarithmic transform. Domain transformation can be confusing because of the difficulty of keeping straight the similar-looking symbols, '$n$' versus '$N$' and '$t$' versus '$T$'. Of course, these symbols are mnemonically chosen. When properly used, these conventions reduce clutter in our formulas. But if they are confusing, you can always fall back to the use of the explicit transformation functions such as $\tau$.

Exercises

**Exercise 8.1:** The text gave the solution $T(N) = 2N - 2^b$ for the recurrence (104). Choose a different DIC to obtain $T(N) = 2N$. ♢

**Exercise 8.2:** Solve recurrence (105) in these cases:
(a) $F(N) = N^k$.
(b) $F(N) = \log N$. ♢

**Exercise 8.3:** (a) Solve the following four recurrences using domain transformation:

$$T(N) = T(N/2) + \begin{cases} \lg N & (i) \\ 1 & (ii) \\ 1/\lg N & (iii) \\ 1/\lg^2 N & (iv) \end{cases}.$$

(b) Generalize the above result: solve the recurrence $T(N) = T(N/2) + \lg^c N$ for all real values of $c$. ♢

**Exercise 8.4:** Justify the simplification step (iv) in §1 (where we replace $\lceil n/2 \rceil$ by $n/2$). ♢

**Exercise 8.5:** When does $T_0(N) = T_0(N/2) + F(N)$ and $T_1(N) = T_1(3 + N/2) + F(N)$ have the same $\Theta$-order? ♢

**Exercise 8.6:** Construct examples where you need to compose two or more of the above domain transformations. ♢

End Exercises

## §9. Range Transformation

**¶38.** A transformation of the range is sometimes called for. For instance, consider

$$T(n) = 2T(n-1) + n.$$

To put this into standard form, we could define

$$t(n) := \frac{T(n)}{2^n}$$

and get the standard form recurrence

$$t(n) = t(n-1) + \frac{n}{2^n}.$$

Telescoping gives us a series of the type in equation (66), which we know how to sum. Specifically, $t(n) = \sum_{x \geqslant 1}^{n} \frac{x}{2^x} = \Theta(1)$ as $f(x) = x/2^x$ is exponentially decreasing. Hence $T(n) = \Theta(2^n)$.

We have transformed the range of $T(n)$ by introducing a multiplicative factor $2^n$: this factor is called the **summation factor**. The reader familiar with linear differential equations will see an analogy with "integrating factor". (In the same spirit, the previous trick of domain transformation is simply a "change of variable".)

In general, a range transformation converts a recurrence of the form

$$T(n) = c_n T(n-1) + F(n) \tag{111}$$

into standard form. Here $c_n$ is a constant depending on $n$. Let us discover which summation factor will work. If $C(n)$ is the summation factor, we get

$$t(n) := \frac{T(n)}{C(n)},$$

and hence

$$\begin{aligned}
t(n) &= \frac{T(n)}{C(n)} \\
&= \frac{c_n}{C(n)} T(n-1) + \frac{F(n)}{C(n)} \\
&= \frac{T(n-1)}{C(n-1)} + \frac{F(n)}{C(n)}, \qquad \text{(provided } C(n) = c_n C(n-1)\text{)} \\
&= t(n-1) + \frac{F(n)}{C(n)}.
\end{aligned}$$

Thus we need $C(n) = c_n C(n-1)$ which expands into

$$C(n) = c_n c_{n-1} \cdots c_1.$$

Exercises

**Exercise 9.1:** (a) Reduce the following recurrence

$$T(n) = 4T(n/2) + \frac{n^2}{\lg n}$$

to standard form. Solve it exactly when $n$ is a power of 2.
(b) Extend the solution of part(a) to general $n$ using our generalized Harmonic numbers $H^{(-1)}(n)$ for real $n \geqslant 2$. State your DIC explicitly. ♢

**Exercise 9.2:** Repeat the previous question for the following recurrences:
(a) $T(n) = 4T(n/2) + \frac{n^2}{\lg^2 n}$
(b) $T(n) = 4T(n/2) + \frac{n^2}{\sqrt{\lg n}}$. ♢

**Exercise 9.3:** Solve the recurrence $T(n) = 5T(n-1) + f(n)$ for $f(n) = 1$, $f(n) = 5^n$ and $f(n) = 10^n$ (respectively) ♢

**Exercise 9.4:** Z.H. proposed to transform the recurrence $T(n) = 100T(n-1) + f(n)$ by using range transformation $t(n) = T(n)/100$. Convince Z.H. that this is futile. ♢

**Exercise 9.5:** Use the EGVS Method to solve the following recurrences
(a) $T(n) = n + 8T(n/2)$.
(b) $T(n) = n + 16T(n/4)$.
(c) Can you generalize your results in (a) and (b) to recurrences of the form $T(n) = n + aT(n/b)$ when $a, b$ are in some special relation? ♢

**Exercise 9.6:** Solve the recurrence (111) in the case where $c_n = 1/n$ and $F(n) = 1$. ♢

**Exercise 9.7:** Solve $T(N) = 100T(N/10) + N^2/\sqrt{\log N}$ using transformations. Assume $\log N$ is to the base 10. ♢

**Exercise 9.8:** Consider the following recurrence

$$T(n) = n^c + aT(n/b)$$

where $a > 0$ and $b > 1$.
(i) Use the Rote (EGVS) method; you must clearly indicate each of the 4 stages (E, G, V, and S) to obtain an open sum.
(ii) Deduce the $\Theta$-order of $T(n)$ depending on the nature of the real constants $a, b, c$. ♢

End Exercises

## §10. The Master Theorem

We first look at a recurrence that does fall under our transformation techniques: the **master recurrence** is

$$T(n) = aT(n/b) + f(n) \tag{112}$$

where $a > 0, b > 1$ are real constants and $f(n)$ is the "forcing" (or driving) function. Our goal is to prove the so-called **Master Theorem** which provides a "cookbook" formula for solutions of the master recurrence. There is a critical constant $w := \log_b a$ that we call the **watershed exponent** for the Master Recurrence.

*one highlight of this chapter!*

**Theorem 10 (Master Theorem)** *The master recurrence (112) has solution:*

$$T(n) = \Theta\begin{cases} n^w, & \text{if } f(n) = \mathcal{O}(n^{w-\epsilon}), \text{ for some } \epsilon > 0, & \text{CASE}(-) \\ n^w \log n, & \text{if } f(n) = \Theta(n^w), & \text{CASE}(0) \\ f(n), & \text{if } af(n/b) \leqslant cf(n) \text{ for some } c < 1. & \text{CASE}(+) \end{cases}$$

We have already seen several instances of this theorem. The solution to the mergesort recurrence $T(n) = 2T(n/2) + n$ falls under CASE(0) of this theorem. Another famous one is Strassen's 1969 algorithm for multiplying two $n \times n$ matrices in subcubic time. Strassen's recurrence $T(n) = 7T(n/2) + n^2$, has solution $T(n) = \Theta(n^{\lg 7})$ which falls under CASE($-$).

Evidently, the Master Recurrence is the recurrence to solve if we manage to solve a problem of size $n$ by breaking it up into $a$ subproblems each of size $n/b$, and merging these $a$ sub-solutions in time $f(n)$. The recurrence was systematically studied by Bentley, Haken and Saxe [1]. Solving it requires a combination of domain and range transformation. Our real formulation has greatly generalized the original setting of the Master Recurrence.

Finally, it may be noted that the 3 cases of the Master Theorem is intimately connected to our Trichotomy Theorem (Theorem 9) for generalized Harmonic numbers.

**¶39. Proof of the Master Theorem.** First apply a domain transformation by defining a new function $t(k)$ from $T(n)$, where $k = \log_b(n)$:

$$t(k) := T(b^k) \quad (\text{for all } k \in \mathbb{R}). \tag{113}$$

Then (112) transforms into

$$t(k) = a\,t(k-1) + f(b^k).$$

Next, transform the range by using the summation factor $1/a^k$. This defines a function $s(k)$ from $t(k)$:

$$s(k) := t(k)/a^k. \tag{114}$$

Now $s(k)$ satisfies a recurrence in standard form:

$$\begin{aligned} s(k) = \frac{t(k)}{a^k} &= \frac{t(k-1)}{a^{k-1}} + \frac{f(b^k)}{a^k} \\ &= s(k-1) + \frac{f(b^k)}{a^k} \end{aligned}$$

Telescoping, we get

$$s(k) = s(\{k\}) + \sum_{i \geqslant 1}^{k} \frac{f(b^i)}{a^i} = \sum_{i \geqslant 1}^{k} \frac{f(b^i)}{a^i}. \tag{115}$$

where $\{k\}$ is the fractional part of $k$ (recall that $k$ is real), and by DIC, we chose $s(x) = 0$ for $x < 1$. We now back substitute this solution to determine the solution in terms of the original function $T(n)$:

$$\begin{aligned} T(n) &= t(\log_b n) && (\text{by } (113)) \\ &= a^{\log_b n} s(\log_b n) && (\text{by } (114)) \\ &= n^{\log_b a} \sum_{i \geqslant 1}^{\log_b n} \frac{f(b^i)}{a^i}. && (\text{by } (115)) \end{aligned} \tag{116}$$

This is the general solution to the master recurrence. Notice that our derivation is completely rigorous (it works for all real $a, b, n$). But $T(n)$ is expressed as an open sum, and we need a closed formula. *Now, we cannot proceed further without knowing the nature of the function $f$.*

We need another important insight. Let us call the function

$$W(n) = n^{\log_b a} = n^w \tag{117}$$

the **watershed function** for our recurrence. The Master Theorem considers three cases for $f$. These cases are obtained by comparing $f$ to $W(n)$. The easiest case is where $f$ and $W$ have the same $\Theta$-order (CASE(0)). The other two cases are where $f$ grows "polynomially slower" (CASE($-$)) or "polynomially faster" (CASE($+$)) than the watershed function.

*The Master Theorem is a trichotomy!*

**CASE**(0) This is when $f(n)$ satisfies

$$f(n) = \Theta(n^{\log_b a}) = \Theta(a^{\log_b n}). \tag{118}$$

Then $f(b^i) = \Theta(a^i)$ and plugging into (116), we get $T(n) = n^w \sum_{i\geqslant 1}^{\log_b n} \Theta(1) = \Theta(n^w \log n)$.

**CASE**($-$) This is when $f(n)$ **grows polynomially slower** than the watershed function:

$$f(n) = \mathcal{O}(n^{-\epsilon+\log_b a}), \tag{119}$$

for some $\epsilon > 0$. Then $f(b^i) = \mathcal{O}(b^{i(\log_b a-\epsilon)}) = \mathcal{O}(a^i b^{-i\epsilon})$. Plugging into (116), we get $T(n) = n^w \sum_{i\geqslant 1}^{\log_b n} \mathcal{O}(b^{-i\epsilon}) = \Theta(n^w)$ since $b^{-\epsilon} < 1$ implies the summation is decreasing exponentially.

**CASE**($+$) This is when $f(n)$ satisfies the **regularity condition**

$$af(n/b) \leqslant cf(n) \text{ (ev.)} \tag{120}$$

for some $c < 1$. Expanding this,

$$\begin{aligned} f(n) &\geqslant \frac{a}{c} f\left(\frac{n}{b}\right) \geqslant \frac{a^2}{c^2} f\left(\frac{n}{b^2}\right) \geqslant \cdots \\ &\geqslant \left(\frac{a}{c}\right)^{\lfloor \log_b n\rfloor} f(C) \\ &= \Omega(n^{\epsilon+\log_b a}), \end{aligned}$$

where $\epsilon = -\log_b c > 0$, and $C = n/b^{\lfloor \log_b n\rfloor}$. We have just proven that the regularity condition implies that $f(n)$ **grows polynomially faster** than the watershed function:

$$f(n) = \Omega(n^{\epsilon+\log_b a}). \tag{121}$$

Writing $k = \log_b n$, we see that regularity implies $f(b^{k-i}) \leqslant (c/a)^i f(b^k)$. Plugging into (116),

$$\begin{array}{rcll} T(n) = n^w \sum_{i\geqslant 1}^{k} f(b^i)/a^i & = & n^w \sum_{i=0}^{\lfloor k-1\rfloor} f(b^{k-i})/a^{k-i} & \text{(switch to ascending sum)} \\ & \leqslant & n^w \sum_{i=0}^{\lfloor k-1\rfloor} (c/a)^i f(b^k)/a^{k-i} & \text{(by regularity of } f\text{)} \\ & = & f(b^k) \sum_{i=0}^{\lfloor k-1\rfloor} c^i & \text{(since } n^w = a^k\text{)} \\ & = & \mathcal{O}(f(b^k)) = \mathcal{O}(f(n)). & \end{array}$$

Since $T(n) \geqslant f(n)$, we have shown that $T(n) = \Theta(f(n)$.

This concludes our proof of the Master Theorem (Theorem 10).

**¶40. Uses of the Master Theorem.** Informally, we describe CASE($+$) as the case when the forcing function $f(n)$ is polynomially faster than $W(n)$. But the actual requirement is somewhat stronger, namely the regularity condition (120). In applications of the Master Theorem, this case is usually the least convenient to check.

We can take advantage of the fact that checking if a function $f(n)$ is polynomially faster (or slower) than $W(n)$ is usually easier to check (just by "inspection"). Hence we normally begin by first verifying the polynomially faster condition, equation (121). If so, we then check the stronger regularity condition (120). To illustrate this process, consider the recurrence

$$T(n) = 3T(n/10) + \sqrt{n}/\lg n.$$

We note that $\alpha = \log_{10} 3 < \log_9 3 = 1/2$ and so $n^{\alpha+\epsilon} \leqslant \sqrt{n}/\lg n$ (ev.), confirming equation (121). We now suspect that CASE($+$) holds, and must verify that

$$cf(n) \geqslant 3f(n/10) \tag{122}$$

for some $0 < c < 1$. This holds, provided

$$\begin{array}{rcl} c^{\frac{\sqrt{n}}{\lg n}} & \geqslant & 3^{\frac{\sqrt{n/10}}{\lg(n/10)}} \\ \Leftarrow\quad c & \geqslant & \sqrt{9/10}^{\frac{\lg n}{\lg(n/10)}}. \end{array}$$

Since $(\lg n)/(\lg(n/10)) \to 1$ as $n \to \infty$, it is sufficient to choose any $c$ satisfying $1 > c > \sqrt{9/10}$.

The polynomial version of the theorem is perhaps most useful:

**Corollary 11** *Let $a > 0, b > 1$ and $k$ be constants. The solution to $T(n) = aT(n/b) + n^k$ is given by*

$$T(n) = \Theta \begin{cases} n^{\log_b a}, & \text{if } \log_b a > k \\ n^k, & \text{if } \log_b a < k \\ n^k \lg n, & \text{if } \log_b a = k \end{cases}$$

What if the values $a, b$ in the master recurrence are not constants but depends on $n$? For instance, attempting to apply this theorem to the recurrence

$$T(n) = 2^n T(n/2) + n^n$$

(with $a = 2^n$ and $b = 2$), we obtain the false conclusion that $T(n) = \Theta(n^n \log n)$. See Exercises. The paper [18] treats the case $T(n) = a(n)T(b(n)) + f(n)$. For other generalizations of the master recurrence, see [17].

**¶41. Graphic Interpretation of the Master Recurrence.** The expansion of the Master Recurrence is commonly shown as a recursion tree with branching factor of $a$ at each internal node, and every leaf of the tree is at level $\log_b a$. This representation associates a "size" of $n/b^i$ and "cost" of $f(n/b^i)$ to each node at level $i$ (root is at level $i = 0$). Then $T(n)$ is just the sum of the costs at all the nodes. The Master Theorem says this: In case (0), the total cost associated with nodes at any level is $\Theta(n^{\log_b a})$ and there are $\log_b n$ levels giving an overall cost of $\Theta(n^{\log_b a} \log n)$. In case (+1), the cost associated with the root is $\Theta(T(n))$, since the cost of the root is $f(n)$. In case (−1), the total cost associated with the leaves is $\Theta(T(n))$: there are $n^{\log_b a} = a^{\log_b n}$ leaves and if each leaf has unit cost, these costs sum up to $\Theta(T(n))$. Of course, this "recursion tree" is not well-defined unless $a$ and $\log_b a$ are integers. So we should remember this is only a useful mnemonic for how the Master Theorem works.

*Draw the recursion tree with a grain of salt!*

**¶42. Beyond the Master Theorem.** Time to make a confession: this section is located deep in this Chapter. In reality, we could have proven the Master Theorem using a direct argument, after we introduced Basic Sums in §6 . But the detour through summation techniques, domain and range transformations has its value: it would allow us to obtain tight bounds even when the forcing function has forms such as $n \log n$ or $n^2/\log n$.

Indeed, several authors[18] have extended the Master Theorem to forcing functions of the form $f(n) = n^k \log^c n$ for all $k, c \in \mathbb{R}$. If $k$ is not equal to the watershed constant, we already know the answer from the Master Theorem. So the interesting case is when $k = \log_b a$. Then there are four possible cases:

*no more trichotomy!*

[18] Unfortunately, the analysis is sometimes partial.

**Theorem 12 (Extended Master Theorem)** *Assume $f(n)$ be the forcing function of the Master Recurrence, and $W(n) = n^{\log_b a}$ is the watershed function. Then the solution to the master recurrence is*

$$T(n) = \Theta \begin{cases} f(n) & \text{if } f(n) \text{ satisfies the regularity condition.} & \text{CASE}(+) \\ W(n)\log^{c+1} n & \text{if } f(n) = \Theta(W(n)\log^c n), \quad c > -1, & \text{CASE}(0) \\ W(n)\log\log n & \text{if } f(n) = \Theta(W(n)\log^c n), \quad c = -1, & \text{CASE}(1) \\ W(n) & \text{if } f = O(W(n)\log^c n), \quad c < -1, & \text{CASE}(-) \end{cases}$$

It is instructive to compare the Extended Master Theorem (EMT) with the Master Theorem (MT):

- CASE(+) of MT is identical to CASE(+) of EMT.
- CASE(0) of MT follows from CASE(0) of EMT (just let $c = 0$).
- Likewise, CASE(−) of MT follows from CASE(−) of EMT.
- But CASE(1) of EMT has no analogue in MT.

But even this generalization does not capture the recurrence that comes from the Schönhage-Strassen recurrence for integer multiplication. For this, we need further generalizations. The idea is to consider $f(n)$ to be any product of powers of iterated logarithms (which we call **EL-functions**). The "ultimate" generalization is proved in [19], with infinitely many cases (CASE(0) and CASE(1) are just the first two instances).

In the next section, however, we consider generalizations of a different nature – we look at a generalization of the Master Recurrence itself.

Exercises

**Exercise 10.1:** Which is the faster growing function: $T_1(n)$ or $T_2(n)$ where

$$T_1(n) = 6T_1(n/2) + n^3, \qquad T_2(n) = 8T_2(n/2) + n^2.$$

◇

**Exercise 10.2:** Suppose $T(n) = n + 3T(n/2) + 2T(n/3)$. Joe claims that $T(n) = O(n)$, Jane claims that $T(n) = O(n^2)$, John claims that $T(n) = O(n^3)$. Who is closest to the truth?

Do not use the Multiterm Master Theorem. You may appeal to the Master Theorem or real induction. ◇

**Exercise 10.3:** Use the Master Theorem to solve the following recurrences arising from matrix multiplication. Be sure to justify the case you choose.

(a) It is easy to see how to recursively multiply two $n \times n$ matrices asymptotically $T(n) = 8T(n/2) + n^2$ time:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{bmatrix}$$

What is the solution $T(n)$ using Master theorem?

(b) Strassen (1969) showed that you can actually save one sub-matrix multiplication, giving the recurrence $S(n) = 7S(n/2) + n^2$. Use the Master theorem to determine $S(n)$.

(c) Virginia Vassilevska Williams shown that we can multiply $n \times n$ matrices in time $T(n) = \mathcal{O}(n^{2.3727})$ operations: "Multiplying matrices faster than Coppersmith-Winograd" (STOC 887–898). Because of such a result, we call 2.3727 a **matrix exponent**. The previous record for the matrix exponent is 2.376. The hidden constants in such algorithms render them impractical at the present time. Nevertheless, it is of theoretical interest to obtain smaller matrix exponents. It is known that $T(n) = \Omega(n^2)$, so a matrix exponent must be at least 2. It has been conjectured that there are exponents arbitrarily close to 2. Suppose you read in Scientific American that someone has discovered a marvelous way of multiplying $2 \times 2$ matrices using only $a$ multiplications, and the recurrence $T(n) = aT(n/2) + n^2$, improving on the 2.3727 exponent. What is the largest possible value of $a$? What do you think is the likelihood of such a result? ♢

*"State" means: no proofs needed*

**Exercise 10.4:** State the Θ-order solution to the following recurrences:

$$\begin{aligned} T(n) &= 10T(n/10) + \log^{10} n. \\ T(n) &= 100T(n/10) + n^{10}. \\ T(n) &= 10T(n/100) + (\log n)^{\log\log n}. \\ T(n) &= 16T(n/4) + 4^{\lg n}. \end{aligned}$$

♢

**Exercise 10.5:** Our proof of the Master Recurrence $T(n) = aT(n/b) + f(n)$ begins with a domain transformation, followed by a range transformation. Can we reverse the order of these 2 transformations? We want to do a range transformation, followed by a domain transformation. What are these transformations? ♢

**Exercise 10.6:** Solve the following using the Master's theorem.
(a) $T(n) = 3T(n/25) + \log^3 n$
(b) $T(n) = 25T(n/3) + (n/\log n)^3$
(c) $T(n) = T(\sqrt{n}) + n$.
HINT: in the third problem, the Master theorem is applicable after a simple transformation. ♢

**Exercise 10.7:** Sometimes the Master Theorem is not applicable directly. But it can still be used to yield useful information. Use the Master Theorem to give as tight an upper and lower bound you can for the following recurrences:
(a) $T(n) = n^3 \log^3 n + 8T(n/2)$
(b) $T(n) = n^2/\log\log n + 9T(n/3)$
(c) $T(n) = 4T(n/2) + 3T(n/3) + n$. ♢

**Exercise 10.8:** Jack invented an algorithm that satisfies the recurrence

$$T_K(n) = aT_K(n/2) + n^\alpha$$

for some constant $a > 1$ and $\alpha$. Jill studied Jack's algorithm and discovered that she could get a faster method with recurrence

$$T_L(n) = (a-1)T_L(n/2) + n^\alpha.$$

What does that tell you about the exponent $\alpha$? NOTE: Jill's improvement is reminiscent of how Karatsuba improved the classical multiplication algorithm. ♢

**Exercise 10.9:** Jack invented an algorithm that satisfies the recurrence

$$T_K(n) = aT_K(n/2) + n^e$$

for some constant $a > 1$. Jill studied his algorithm and discovered that she could get a faster method with recurrence

$$T_L(n) = (a-1)T_L(n/2) + n^e.$$

What does that tell you about the exponent $e$? NOTE: Jill's improvement is reminiscent of how Karatsuba improved the classical multiplication algorithm. ♢

**Exercise 10.10:** We want to improve on Karatsuba's multiplication algorithm. We managed to subdivide a problem of size $n$ into $a \geqslant 2$ subproblems of size $n/4$. After solving these $a$ subproblems, we could combine their solutions in $O(n)$ time to get the solution to the original problem of size $n$. To beat Karatsuba, what is the maximum value $a$ can have? $\diamondsuit$

**Exercise 10.11:** Suppose algorithm $A_1$ has running time satisfying the recurrence

$$T_1(n) = aT(n/2) + n$$

and algorithm $A_2$ has running time satisfying the recurrence

$$T_2(n) = 2aT(n/4) + n.$$

Here, $a > 0$ is a parameter which the designer of the algorithm can choose. Compare these two running times for various values of $a$. $\diamondsuit$

**Exercise 10.12:** Say whether $T_1(n) \lll T_2(n)$ or $T_1(n) \ggg T_2(n)$ where

$$T_1(n) = 8T_1(n/4) + n^{1.5}, \qquad T_2(n) = 6T_2(n/3) + n^2.$$

Briefly justify using Master Theorem; do not use calculators. $\diamondsuit$

**Exercise 10.13:** Suppose

$$T_0(n) = 18T_0(n/6) + n^{1.5}$$

and

$$T_1(n) = 32T_1(n/8) + n^{1.5}.$$

Which is the correct relation: $T_0(n) = \Omega(T_1(n))$ or $T_0(n) = \mathcal{O}(T_1(n))$? Do this exercise without using a calculator or its equivalent; instead, use inequalities such as $\log_8(x) < \log_6(x)$ (for $x > 1$) and $\log_6(2) < 1/2$. $\diamondsuit$

**Exercise 10.14:** Solve the master recurrence when $f(n) = n^{\log_b a} \log^k n$, for all $k \in \mathbb{R}$. You need to use the transformation methods in order to determine the $\Theta$-order correctly. (Be careful when $k = -1$.) $\diamondsuit$

**Exercise 10.15:** Show that the master theorem applies to the following variation of the master recurrence:

$$T(n) = a \cdot T(\frac{n+c}{b}) + f(n)$$

where $a > 0, b > 1$ and $c$ is arbitrary. $\diamondsuit$

**Exercise 10.16:**

(a) Solve $T(n) = 2^n T(n/2) + n^n$ by direct expansion.

(b) To what extent can you generalize the Master theorem to handle some cases of $T(n) = a_n T(n/b_n) + f(n)$ where $a_n, b_n$ are both functions of $n$? $\diamondsuit$

**Exercise 10.17:** Let $W(n)$ be the watershed function of the master recurrence. In what sense is the "watershed function" of the next order equal to $W(n)/\ln n$? $\diamondsuit$

**Exercise 10.18:**

(a) Let

$$s(n) = \sum_{i \geqslant 1}^{n} \frac{\lg i}{i}$$

Prove that $s(n) = \Theta(\lg^2 n)$ directly. Note that our theory of growth types is no help here. For the lower bound, you can use real induction, and the fact that for $n \geqslant 2$, we have

$$\ln(n) - (2/n) < \ln(n-1) < (\ln n) - (1/n).$$

(b) Using the domain/range transformations to solve the following recurrence:

$$T(n) = 2T(n/2) + n\frac{\lg\lg n}{\lg n}.$$

◇

**Exercise 10.19:** Consider the recurrence $T(n) = aT(n/b) + \frac{n^4}{\log n}$ where $a > 0$ and $b > 1$. Describe the set $S$ of all pairs $(a, b)$ for which the Master Theorem gives a solution for this recurrence. Do not describe the solutions. You must describe the set $S$ in the simplest possible terms. ◇

**Exercise 10.20:** (Reif and Sen, 1988)The following recurrences arises in the analysis of a parallel algorithm for hidden-surface removal:

$$T(n) = T(2n/3) + \lg n \lg\lg n$$

Another version of the algorithm [18] leads to

$$T(n) = T(2n/3) + (\lg n)/\lg\lg n.$$

Solve for $T(n)$ in both cases. ◇

End Exercises

## §11. The Multiterm Master Theorem

The Master recurrence (112) can be generalized to the following **multiterm master recurrence**:

$$T(n) = d(n) + \sum_{i=1}^{k} a_i T\left(\frac{n}{b_i}\right) \tag{123}$$

where $k \geqslant 1$, $a_i > 0$ (for all $i = 1, \ldots, k$) and $b_1 > b_2 > \cdots > b_k > 1$. Again, $d(n)$ is the driving (or forcing) function. When $k = 2$, we have the following examples of 2-term master recurrences:

$$T(n) = T(c_1 n) + T(c_2 n) + n, \qquad (c_1 + c_2 < 1). \tag{124}$$

$$T(n) = T(n/2) + T(n/4) + 1. \tag{125}$$

The first recurrence (124) arise in linear time selection algorithms (see Chapter XI). There are many versions of this algorithm with different choices for the constants $c_1, c_2$. E.g., $c_1 = 7/10, c_2 = 1/5$. The second recurrence (125) arose in Computational Geometry in a data structure called **conjugation tree**. Edelsbrunner and Welzl [3] introduced the structure for solving the **point retrieval problem**.

**¶* 43. Conjugation Tree.** This problem here illustrates how a multiterm recurrence arises in Computational Geometry. Consider the **Point Retrieval Problem**: given a set $S$ of $n$ distinct points in the plane, we want to construct a data structure $D(S)$ to store $S$, so that we can subsequently use $D(S)$ to efficiently answer **half-plane queries** of the following form: *given a half plane $H$, return the set $H \cap S$.* Assume $H$ is the open half-plane $\ell^+$ (resp., $\ell^-$) that lies to the left (resp., right) of a directed line

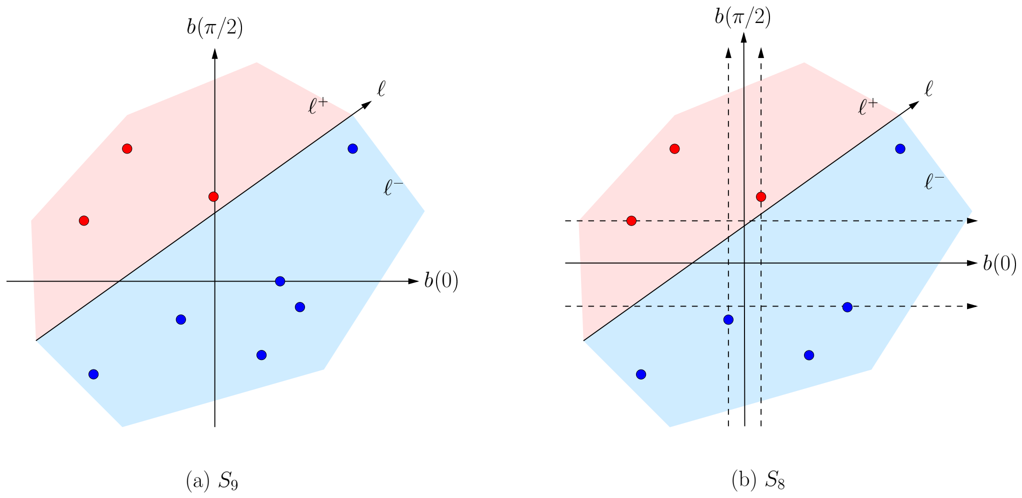


Figure 7: Half-plane query on $S_9$ and $S_8$

$\ell$. A directed line $\ell$ can be represented by any pair of distinct points $p, q$ on the line, with the direction going from $p$ to $q$. Alternatively, $\ell$ can be represented by a linear equation $L(x, y) = ax + by + c = 0$ and the two half-planes can have equations $L(x, y) > 0$ and $L(x, y) < 0$, respectively. In Figure 7(a), if $S_9$ is a set of 9 points centered at the small discs and $H = \ell^+$ (resp. $H = \ell^-$) then we should return the set of three red (resp., six blue) points. We describe a lovely construction for $D(S)$ from Edelsbrunner and Welzl (1986).

The line $\ell$ is a **bisector** of $S$ if $|\ell^+ \cap S| \leqslant n/2$ and $|\ell^- \cap S| \leqslant n/2$. Suppose $|S| = n$ is odd. It is easy to see that any bisector of $S$ must pass through at least one point of $S$, and this bisector is unique for any orientation $\theta \in [0, 2\pi)$. Let $\beta(S, \theta)$ (or $\beta(\theta)$ if $S$ is understood) denote this unique bisector. For instance, if $\theta = 0$ then $\beta(\theta)$ is horizontal, pointing to the right. In Figure 7, we display the two bisectors $\beta(0)$ and $\beta(\pi/2)$. Suppose $n$ even. If there is more than one bisector of $S$ with orientation $\theta$, then we see that there is a strip of the plane bounded by two lines with orientation $\theta$ such that every line lying in this strip is a bisector. Let $\beta(S, \theta)$ denote the line that lies in the middle of this strip. In Figure 7(b), $S_8$ is obtained from $S_9$ by deleting one point. There are infinitely many horizontal bisectors of $S_8$, lying in the strip between the two dashed horizontal lines, and $\beta(S, 0)$ is the line in the middle of this strip. Therefore, for every $\theta$ and any $S$, we have defined a unique bisector denoted $\beta(\theta, S)$. If $\beta(\theta, S)$ contains two or more points of $S$, we call $\theta$ an **exceptional orientation** of $S$.

The set of angles has the topology of a circle. Given $\theta_1$ and $\theta_2$ where $0 \leqslant \theta_1 < \theta_2 < 2\pi$, we define two kinds of **circular intervals**:

$$\begin{aligned} (\theta_1, \theta_2) &:= \{\theta : \theta_1 < \theta < \theta_2\} \\ (\theta_2, \theta_1) &:= \{\theta : \theta_2 < \theta < 2\pi \text{ or } 0 \leqslant \theta < \theta_2\}. \end{aligned}$$

The endpoints of these intervals are $\theta_1$ and $\theta_2$. A circular interval $J \subseteq [0, 2\pi)$ is said to be **exceptional** if its endpoints are exceptional, but no angle in $J$ is exceptional.

**Lemma 13** *Let $J \subseteq [0, 2\pi)$ be a circular interval. If $J$ is exceptional, then there exists a unique "pivot" $p$ such that for all $\theta \in J$, the line $\beta(S, \theta)$ passes through $p$. Furthermore:*
*(i) If $|S|$ is odd, then $p$ is a point in $S$.*
*(ii) If $|S|$ is even, then there exists a pair of distinct points $p_1, p_2 \in S$ such that $p = (p_1 + p_2)/2$.*

We leave the easy proof to the reader. It follows that the pivots of each exceptional interval of $S$ lies inside the convex hull of $S$.

Consider the pair $(S, b)$ where $b$ is a bisector of $S$. Call $\ell$ a **conjugate bisector** of $(S, b)$ if $\ell$ is a simultaneous bisector of $S \cap b^+$ and of $S \cap b^-$. *Thus, the pair $(b, \ell)$ of lines partitions the plane into 4 (skew) quadrants, each containing at most $|S|/4$ points.* Such an $\ell$ is also[19] known as a **ham sandwich cut** of $S \cap b^+$ and $S \cap b^-$.

**Lemma 14 (Willard)** *Willard Conjugate bisectors exist for any $(S, b)$.*

*Proof.* Let $A = S \cap b^+$ and $B = S \cap b^-$. If both $A$ and $B$ are empty, then any line can be considered conjugate bisectors of $(S, b)$. If one of the sets $A$ or $B$ is empty, then any bisector of the non-empty set would be a conjugate bisector of $(S, b)$. Therefore assume both $A$ and $B$ are non-empty.

Wlog, let the bisector $b$ of $S$ be the positive $x$-axis, and so $A$ comprise those points in $S$ lying $\underline{a}$bove the $x$-axis and $B$ comprise those points in $S$ lying $\underline{b}$elow the $x$-axis. Write $\beta(\theta)$ for $\beta(A, \theta)$ and define the fraction

$$f(\theta) = \frac{|\beta(\theta)^+ \cap B|}{|\beta(\theta)^- \cap B|}.$$

Note that $\beta(0)$ is a horizontal line parallel to the $x$-axis, with $\beta(0)^+ \cap B$ is empty while $\beta(0)^- \cap B = B$. Therefore $f(0) = 0/|B| = 0$. Conversely, $\beta(\pi)$ is the same line as $\beta(0)$ but with the opposite orientation, and therefore $f(\pi) = |B|/0 = \infty$.

We claim that $f(\theta)$ is non-decreasing as $\theta$ increases. First we show that $f(\theta)$ is non-decreasing as $\theta$ varies inside each exceptional interval $J$. By the previous lemma, there is a pivot $p$ in the convex hull of $A$ such that $\beta(\theta)$ is rotating about $p$ as $\theta$ increases. Since $p$ lies above the $x$-axis (i.e., above $b$), we see that $c = |\beta(\theta)^+ \cap B|$ can only increase and $d = |\beta(\theta)^- \cap B|$ can only decrease as $\theta$ increases. Thus $f(\theta) = c/d$ can only increase. Let us fix $\theta$ and let $\Delta = |\beta(\theta) \cap B|$. We have two possibilities: (i) If $\Delta = 0$, then $f(x)$ is the constant $c/d$ for all $x$ ranging in some interval $(\theta - \epsilon, \theta + \epsilon) \subseteq J$ where $\epsilon > 0$. (ii) If $\Delta \geqslant 1$, then

$$f(x) = \begin{cases} \frac{c}{d+\Delta} & \theta - \epsilon < x < \theta, \\ \frac{c}{d} & x = \theta, \\ \frac{c+\Delta}{d} & \theta < x < \theta + \epsilon. \end{cases}$$

It follows that $f(x)$ is a non-decreasing piecewise constant function. If $f(x) = 1$ for some $x$, then we are done. Otherwise, there is some $\theta$ such that $\frac{c}{d+\Delta} < 1 < \frac{c+\Delta}{d}$. This implies $c - \Delta < d < c + \Delta$. Note that $|B| = c + d + \Delta$. Therefore $c/|B| = c/(c + d + \Delta) < c/(2c) = 1/2$. Likewise, $d/|B| = d/(c + d + \Delta) < d/(2d) = 1/2$. This proves $\beta(\theta)$ is a bisector of $B$.

What if $\theta$ is an exceptional orientation for $A$? In this case, $\beta(x)$ has one pivot for $x \in (\theta - \epsilon, \theta)$ and another pivot for $x \in (\theta, \theta + \epsilon)$. But the same argument applies to show that $f(x)$ is non-decreasing at $x = \theta$. Moreover, if $f(\theta - \epsilon) = \frac{c}{d+\Delta} < 1 < \frac{c+\Delta}{d} = f(\theta + \epsilon)$, then $\beta(\theta)$ is a bisector of $B$. **Q.E.D.**

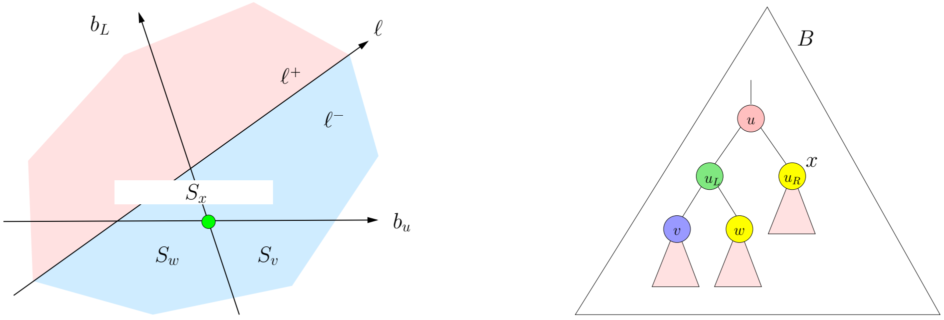
A **conjugation tree** for a set $S$ is a full binary tree $T(B)$ such that

- Each node $u$ of $T(B)$ may be identified with a pair of the form $u = (S_u, b_u)$ where $S_u \subseteq S$ and $b_u$ is a bisector of $S_u$. Call $S_u$ and $b_u$ the **underlying set** and **bisector** at $u$.
- The underlying set at the root is $S$, and the underlying set at any leaf is a singleton.
- Suppose $u = (S_u, b_u)$ is an internal node with children $u_L = (S_L, b_L)$ and $u_R = (S_R, b_R)$. Then $b_L = b_R$ and $b_L$ is a conjugate bisector of $(S_u, b_u)$. Moreover, $S_L = S_u \cap b_u^+$ and $S_R = S_u \cap b_u^-$.

[19] Ham sandwich cuts are more general than conjugate bisectors: it is defined for any two sets $A, B$ of points. We treat only the special case whether $A$ and $B$ are of the form $A = \ell^+ \cap S$ and $B = \ell^- \cap S$. These notions extend to any dimension, and point sets can be replaced by continuous distributions.

**How to use the Conjugate Tree.** The **half-space point retrieval** problem (for $S$) is a preprocessing problem (§I.2) that can be solved as follows: in the pre-processing stage, we construct a conjugation tree $B$ for $S$. In the query stage, given any query line $\ell$, we use $B$ to compute the set $S \cap \ell^+$. any given "query" line $\ell$. We now describe the query algorithm

Given $\ell$, the recursive algorithm begins by "visiting" the root of $B$. It will recursively visit certain nodes of $B$ and will "mark" some of these visited nodes. The union of the underlying sets at the marked nodes would constitute the set $\ell^+ \cap S$.



(a) $Q(u) = Q(w) + Q(x)$

(b) Conjugate tree $B$

Figure 8: (a) Conjugation tree at $u$, (b) Bisectors $b_u$ and $b_L = b_R$

We now describe the recursive algorithm: suppose that we are visiting node $u = (S_u, b_u)$ in $B$. See Figure 8(a).
1. If $u$ is a leaf, we may mark $u$ iff the point in $S_u$ is in $\ell^+$, and terminate.)
2. Otherwise, let the children of $u$ be $u_L = (S_L, b_L)$ and $u_R = (S_R, b_R)$ where $b_L = b_R$. If $u_L$ and $u_R$ are leaves, we can also mark the nodes $u_L$ and $u_R$ as needed; again we terminate.
3. Finally assuming $u_L, u_R$ are non-leaves, we consider the four quadrants defined by $(b_u, b_L)$: One of the four quadrants is fully contained in $\ell^+$ or fully contained in $\ell^-$. Let $v$ denote the grandchild of $u$ corresponding to this quadrant. The underlying set $S_v$ at $v$ is either contained in $S \cap \ell^+$ or disjoint from $S \cap \ell^-$. We **mark** the node $v$ iff $S_v \subseteq S \cap \ell^+$. Then we recursive visit the $w$ sibling of $v$ and the uncle $x$ of $v$. If $v$ is a child of $u_L$ then $x = u_R$, otherwise $x = u_L$.

Let $T(n)$ be the running time of our query algorithm to visit a node $u$ where $|S_u| = n$. In the non-terminal case, it will next visit two other nodes $v$ and $x$ where $|S_v| \leqslant n/4$ and $|S_x| \leqslant n/2$. The amount of work to visit $u$ is $O(1)$. It follows that $T(n)$ satisfies the recurrence (125) in our introduction.

### ¶44. Reducing multiterm to single term master recurrences.

Before providing the general solution, let us see how our previous techniques would fare here. First of all, rote expansion seems hopeless, even for a two-term master recurrence. On a more positive note, the method of real induction can provide us with confirmations of guessed upper and lower bounds – we had already seen such examples. The catch is how do we go about guessing these bounds. But here is an interesting method to use the Master Theorem to provide upper and lower bounds. The idea is to convert our multiterm recurrence into a master recurrence: let $a := \sum_{i=1}^{k} a_i$, $b := \min\{b_i : i = 1, \ldots, k\}$, and $c := \max\{b_i : i = 1, \ldots, k\}$. This defines two master recurrences

*The student is invited to expand the 2-term recurrences...*

$$\begin{aligned} U(n) &= d(n) + aU(n/b), && (126)\\ L(n) &= d(n) + aL(n/c). && (127) \end{aligned}$$

Clearly, $T(n) = O(U(n))$ and $T(n) = \Omega(L(n))$. Then the Master Theorem implies the bound

$$T(n) = \begin{cases} \mathcal{O}(d(n)\log n + n^{\log_b a}), \\ \Omega(d(n) + n^{\log_c a}). \end{cases} \qquad (128)$$

Applying this to the conjugation tree recurrence (125), we obtain

$$T(n) = \begin{cases} \mathcal{O}(n), \\ \Omega(\sqrt{n}). \end{cases}$$

But suppose we first expand our recurrence once:

$$\begin{aligned} T(n) &= \boxed{T(n/2)} + T(n/4) + 1 \\ &= \boxed{T(n/4) + T(n/8) + 1} + T(n/4) + 1 \end{aligned}$$

Now the application of (128) yields the strictly sharper bound:

$$T(n) = \begin{cases} \mathcal{O}(n^{\log_4 3}), \\ \Omega(n^{\log_8 3}). \end{cases}$$

It is clear that this trick can be repeated. We remark that the lower bound can sometimes be improved by omitting terms before taking the maximum to form $c$. E.g., for $T(n) = T(n/2)+T(n/3)+T(n/9)+1$, the above scheme yields $T(n) = \Omega(\sqrt{n})$, but if we first drop the term $T(n/9)$, we get the improvement $T(n) = \Omega(n^{\log_3 2})$.

**¶45. Multiterm Generalization of Master Theorem.** To state the multiterm analogue of the Master Theorem, we must generalize two concepts from the Master Theorem:

(a) Associated with the recurrence (123) is the **watershed exponent**, a real number $\alpha$ such that

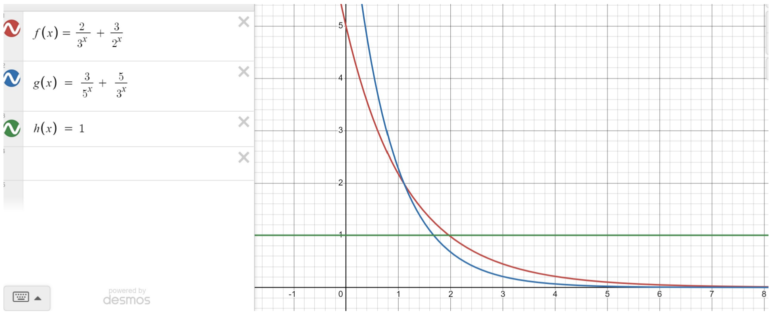
$$\sum_{i=1}^{k} \frac{a_i}{b_i^{\alpha}} = 1. \tag{129}$$

As noted next, there is a unique $\alpha$ that satisfies (129). This is illustrated in Figure 9. As usual, let $W(n) = n^{\alpha}$ denote the watershed function.

(b) The recurrence (123) gives rise to a **generalized regularity condition** on the driving (or forcing) function $d(n)$, namely,

$$\sum_{i=1}^{k} a_i d\Big(\frac{n}{b_i}\Big) \leqslant c d(n) \text{ (ev.)} \tag{130}$$

for some $0 < c < 1$. As for the Master Recurrence, regularity of $d$ implies that $d(n) = \Omega(n^{\alpha+\varepsilon})$ for some $\varepsilon > 0$ (see below).



$f(x) = \frac{2}{3^x} + \frac{3}{2^x}$

$g(x) = \frac{3}{5^x} + \frac{5}{3^x}$

$h(x) = 1$

Figure 9: Functions $f_{red}(x) = \frac{2}{3^x} + \frac{3}{2^x}$ and $f_{blue}(x) = \frac{3}{5^x} + \frac{5}{3^x}$

Let us see why $\alpha$ is uniquely defined by (129): Let us introduce the function $f(x) = \sum_{i=1}^{k} \frac{a_i}{b_i^x}$ corresponding to the Multiterm Master Recurrence (123). It follows that its watershed constant is just $\alpha = f^{-1}(1)$. It is easily seen that $f(x)$ monotone decreasing with $x$, with $f(x) \to 0$ as $\alpha \to \infty$, and $f(x) \to \infty$ as $x \to -\infty$. So $f^{-1}(1)$ exists and is unique. This is illustrated by the plot of two functions in Figure 9: visibly, $f_{red}^{-1}(1) = 2 - \varepsilon$ and $f_{blue}^{-1}(1) = 1.6 + \varepsilon'$.

**Theorem 15 (Multiterm Master Theorem)** *Assume $d(n)$ is a growth function.*

$$T(n) = \Theta \begin{cases} n^\alpha \log n & \text{if } d(n) = \Theta(n^\alpha) & \textit{[CASE(0)]} \\ n^\alpha & \text{if } d(n) = \mathcal{O}(n^{\alpha-\varepsilon}), \textit{for some } \varepsilon > 0, & \textit{[CASE(}-\textit{)]} \\ d(n) & \text{if } d \text{ satisfies the regularity condition (130).} & \textit{[CASE(}+\textit{)]} \end{cases}$$

Recall ¶18 that $d$ is a growth function means that eventually it is eventually defined and non-decreasing, and it is unbounded.

Before proving this result, let us see its application to the conjugation tree recurrence (125). The watershed constant $\alpha$ satisfies the equation $\frac{1}{2^\alpha} + \frac{1}{4^\alpha} = 1$. Writing $x = \frac{1}{2^\alpha}$, we get the equation $x + x^2 = 1$. The positive solution to this quadratic equation is $x = 2^{-\alpha} = (-1+\sqrt{5})/2$. This yields $\alpha = 1 - \lg(-1+\sqrt{5}) \sim 0.695$. Edelsbrunner and Welzl said that they obtained this $\alpha$ by "an analogy with Fibonacci recurrences"; but we now know that it can be systematically derived. They proved that $T(n) = O(n^\alpha)$; our theorem further shows that their bound is $\Theta$-tight.

**¶46. Proof of Multiterm Master Theorem.** We use real induction. Let us write the multiterm recurrence in the form

$$T(n) = G(x, T(g_1(x)), \ldots, T(g_k(x)))$$

where $G(x, t_1, \ldots, t_k) = d(x) + \sum_{i=1}^k a_i t_i$ and $g_i(x) = x/b_i$. Clearly, $G(x, t_1, \ldots, t_k)$, $g_i(x)$ are growth functions; and $d(x)$ is a growth function by assumption. According to Theorem 3, the Real Basis Step is automatic. Thus, we only have to establish the Real Induction Step

CASE(0): Assume that $d(n) = \Theta_1(W(n))$. We will show that $T(n) = \Theta_2(W(n)\log n)$. We have

$$\begin{array}{rcll} T(n) & = & d(n) + \sum_{i=1}^k a_i T\left(\frac{n}{b_i}\right) & \\ & = & \Theta_1(n^\alpha) + \sum_{i=1}^k a_i \Theta_2\left(\left(\frac{n}{b_i}\right)^\alpha \log\left(\frac{n}{b_i}\right)\right) & \text{(by induction)} \\ & = & \Theta_1(n^\alpha) + \Theta_2(n^\alpha)\left[\sum_{i=1}^k \frac{a_i}{b_i^\alpha} \log\left(\frac{n}{b_i}\right)\right] & \\ & = & \Theta_1(n^\alpha) + \Theta_2(n^\alpha)\left[\log n - D\right], & \text{(where } D = \sum_{i=1}^k \frac{a_i}{b_i^\alpha}\log(b_i) \text{ and using (129))} \\ & = & \Theta_2(n^\alpha \log n). & \end{array}$$

Let us elaborate on the last equality. Suppose $d(n) = \Theta_1(n^\alpha)$ amounts to the inequalities $c_1 W(n) \leqslant d(n) \leqslant C_1 W(n)$ (ev.). We must choose $c_2, C_2$ such that $c_2 W(n) \log n \leqslant T(n) \leqslant C_2 W(n) \log n$ (ev.). Note that $D > 1$. We check that the following inequality is sufficient:

$$C_2 \geqslant C_1/D, \qquad c_2 \leqslant c_1/D.$$

CASE($-$): Assume $0 \leqslant d(n) \leqslant D_1 n^{\alpha-\varepsilon}$ for some $\varepsilon > 0$. The lower bound is easy: assume $T(n/b_i) \geqslant c_1(n/b_i)^\alpha$ (ev.) for each $i$. Then[20]

$$\begin{array}{rcll} T(n) & = & d(n) + \sum_{i=1}^k a_i T\left(\frac{n}{b_i}\right) & \\ & \geqslant & \sum_{i=1}^k a_i c_1 (\frac{n}{b_i})^\alpha & \text{(since } d(n) \geqslant 0 \text{ and by induction)} \\ & = & c_1 n^\alpha. & \end{array}$$

The upper bound needs a slightly stronger hypothesis: assume $T(n/b_i) \leqslant C_1 n^\alpha (1 - n^{-\varepsilon})$ (ev.). Then

$$\begin{array}{rcll} T(n) & = & d(n) + \sum_{i=1}^k a_i T\left(\frac{n}{b_i}\right) & \\ & \leqslant & D_1 n^{\alpha-\varepsilon} + \sum_{i=1}^k a_i C_1 \left(\frac{n}{b_i}\right)^\alpha \left[1 - \left(\frac{n}{b_i}\right)^{-\varepsilon}\right] & \text{(by induction)} \\ & = & C_1 n^\alpha - C_1 n^{\alpha-\varepsilon}\left[\sum_{i=1}^k \frac{a_i}{b_i^{\alpha-\varepsilon}} - D_1/C_1\right] & \\ & \leqslant & C_1 n^\alpha - C_1 n^{\alpha-\varepsilon} & \end{array}$$

[20] The fact $d(n) \geqslant 0$ (ev.) is a consequence of "$f \in \mathcal{O}(n^{\alpha-\varepsilon})$" and the definition of the big-Oh notation.

provided $\sum_{i=1}^k a_i/b_i^{\alpha-\varepsilon} \geqslant 1+(D_1/C_1)$. Since $\sum_{i=1}^k a_i/b_i^{\alpha-\varepsilon} > 1$, we can certainly choose a large enough $C_1$ to satisfy this.

CASE(+): The lower bound $T(n) = \Omega(d(n))$ is trivial. As for upper bound, assuming $T(m) \leqslant D_1 d(m)$ (ev.) whenever $m = n/b_i$,

$$\begin{array}{rcll} T(n) & = & d(n) + \sum_{i=1}^k a_i T\left(\frac{n}{b_i}\right) & \\ & \leqslant & d(n) + \sum_{i=1}^k a_i D_1 d(n/b_i) & \text{(by induction)} \\ & = & d(n) + D_1 c d(n) & \text{(by regularity)} \\ & \leqslant & D_1 d(n) & (\text{if } D_1 \geqslant 1/(1-c)) \end{array}$$

This concludes the proof of the Multiterm Master Theorem.

The use of real induction appears to be necessary in this proof: unlike the master recurrence, the multiterm version does not yield to transformations. Again, the generalized regularity condition implies that $d(n) = \Omega(n^{\alpha+\varepsilon})$ for some $\varepsilon > 0$. This is shown by induction:

$$\begin{array}{rcll} d(n) & \geqslant & \frac{1}{c}\sum_{i=1}^k a_i d(n/b_i) & \\ & \geqslant & \frac{1}{c}\sum_{i=1}^k a_i D(n/b_i)^{\alpha+\varepsilon} & (\text{by induction, for some } D > 0) \\ & = & \frac{D}{c} n^{\alpha+\varepsilon} \sum_{i=1}^k \frac{a_i}{b_i^{\alpha+\varepsilon}} & \\ & = & D n^{\alpha+\varepsilon} & (\text{if we choose } c = \sum_{i=1}^k \frac{a_i}{b_i^{\alpha+\varepsilon}}) \end{array}$$

Since $\sum_{i=1}^k \frac{a_i}{b_i^{\alpha}} = 1$, we should be able to choose a $\varepsilon > 0$ to satisfy the last condition. Note that this derivation imposes no condition on $D$, and so $D$ can be determined based on the initial conditions. The above Multiterm Master Theorem in this generality, including an additional fourth case, is first stated and proved in [19].

### Exercises

**Exercise 11.1:** Consider this "3-ary predicate"

$$C(x, y, w) : x^w \leqslant y^w$$

for all $w \in \mathbb{R}$ and $x, y > 0$. Give conditions when $C(x, y, w)$ is true. ♢

**Exercise 11.2:** Let $w$ be the watershed constant for the recurrence

$$T(n) = aT(n/b) + cT(n/d) + 1$$

where $a, c > 0$ and $b, d > 1$.

(a) How do you decide whether $w$ is zero, positive or negative?

(b) Give upper and lower bounds on $w$ using these four parameters:

$$\underline{a} := \min\{a, c\}, \quad \overline{a} := \max\{a, c\}, \tag{131}$$

$$\underline{b} := \min\{b, d\}, \quad \overline{b} := \max\{b, d\}. \tag{132}$$

(c) What are your bounds on $w$ when $(a, b, c, d) = (3, 2, 2, 3)$? Please give the numerical range.

(d) Give an algorithm called Approx$(a, b, c, d, n)$ which returns an approximation of $w$ to $n$ digits of accuracy, i.e., returns a value $\widetilde{w}$ satisfying $|w - \widetilde{w}| < 10^{-n}$. You must write your algorithm using pseudo-code (see ¶I.A.11 in Appendix of Chapter I).

♢

**Exercise 11.3:** Suppose $T(n) = 3T(n/4) + 2T(n/3) + n$. Give upper and lower bounds on $T(n)$ based on the Master (!) Theorem. ♢

**Exercise 11.4:** Suppose $T(n) = n^5 + T(9n/10) + T(17n/20)$.
(a) Use Real Induction to prove that $T(n) = \Theta(n^w)$ for some real $w$. NOTE: do not invoke the Multiterm Master Theorem for this part.
(b) Using a scientific calculator only, determine the $w$ to three significant digits. Explain how you do this calculation.
(c) What is the solution for the related recurrence $T(n) = n^5 + T(9n/10) + T(8n/10)$? You may use the Multiterm Master Theorem for this part. ◇

**Exercise 11.5:** Using the Master Theorem (*not* Multiterm Master Theorem) to provide upper and lower bounds on these recurrence functions. No proofs needed.
(a) State upper and lower bounds on $T(n)$ where

$$T(n) = T(n/2) + T(n/4) + \sqrt{n}.$$

(b) State improved upper and lower bounds over part(a), by first expanding the recurrence *one* step and then invoking Master Theorem. ◇

**Exercise 11.6:** Prove tight upper and lower bounds on $T(n)$ where:
(a) $T(n) = n^3 \log^3 n + 9T(n/3)$.
(b) $T(n) = n^2 \log^3 n + 9T(n/3)$. Using only the Master Theorem (not Multiterm Master Theorem). Be sure to justify the cases used in the Master Theorem. ◇

**Exercise 11.7:** Use the Master Theorem (*not* the Multiterm Master Theorem) to derive a sublinear upper bound on $T(n) = 2T(n/3) + T(n/10) + 1$. Recall some tricks in the text. ◇

**Exercise 11.8:** (What is the recurrence?)
We tell you that the solution to the following recurrence

$$T_2(n) = n + a_1 \cdot T(n/2) + a_2 \cdot T(n/3) + a_3 \cdot T(n/6)$$

has the solution $T_2(n) = \Theta(n \log n)$. What is the recurrence? I.e., determine the constants $a_1, a_2, a_3$. Can you find integer values for the $a_i$'s? ◇

**Exercise 11.9:** Consider the multiterm recurrence $T(n) = T(n/2)+T(n/4)+T(n/8)+1$. Numerically determine watershed constant $\alpha$. Show $\alpha$ up to 3 decimal places. You must describe how you obtain the answer (e.g., perhaps using a hand-calculator). ◇

**Exercise 11.10:** To understand the recurrence $T(n) = T(n/2) + T(n/3) + T(n/4) + n$, we will explore numerically the function $h(x) = 2^{-x}+3^{-x}+4^{-x}$. We want to determine the $\alpha$ such that $h(\alpha) = 1$. For a simple way to do this, use a user-friendly, powerful software like `MATLAB`. For instance, consider the following two lines of `MATLAB` code:

```
>>   h = @(x) 2.ˆ (-x) + 3.ˆ (-x) + 4.ˆ (-x);

>>   for x = 0.9 : 0.1 : 1.2, display([x, h(x)]), end
```

The first line defines the function $h(x)$. The second line is a for-loop where $x$ begins with the value 0.9 and each iteration increases the value of $x$ by 0.1 until $x = 1.2$. Each iteration simply prints the pair $(x, h(x))$ of values. This loop produces the values shown in the first of the following four tables:

| $x$ | $h(x)$ |
|---|---|
| 0.9000 | 1.1951 |
| 1.0000 | 1.0833 |
| 1.1000 | 0.9828 |
| 1.2000 | 0.8923 |

| $x$ | $h(x)$ |
|---|---|
| 1.0700 | 1.0119 |
| 1.0800 | 1.0021 |
| 1.0900 | 0.9924 |
| 1.1000 | 0.9828 |

| $x$ | $h(x)$ |
|---|---|
| 1.0810 | 1.0011 |
| 1.0820 | 1.0001 |
| 1.0830 | 0.9992 |
| 1.0840 | 0.9982 |

| $x$ | $h(x)$ |
|---|---|
| 1.0820 | 1.0001 |
| 1.0821 | 1.0000 |
| 1.0822 | 0.9999 |
| 1.0823 | 0.9998 |

By changing the stepsize and limits of the for-loop, we can get more correct digits with run of the for-loop. Each successive table above is obtained this way, each time giving us an extra digit in the decimal expansion of $\alpha$. Thus, $\alpha \approx 1.0821$. How would you continue this experiment to determine the first 100 digits of $\alpha$? ♢

**Exercise 11.11:** Let $M(n,k)$ be the number of worst case number of comparisons (in the comparison-tree model) to find the rank $k$ element among $n$ elements (for any $k = 1, \ldots, n$). Note that the rank of an element in a set is the number of elements that are greater than or equal to it. When $k = \lceil n/2 \rceil$, we call this the **median problem**. Also, let $M(n) = \max\{M(n,k) : k = 1, \ldots, n\}$.
(i) It can be shown that $M(n) = M(n/5) + M(7n/10) + Cn$ for some constant $C$. Determine the watershed constant $\alpha$ for this recurrence. We suggest you use a pocket calculator and determine $\alpha$ up to 2 digits, using a simple binary search (one digit at a time).
(ii) Conclude from the Multiterm Master Theorem that $M(n) = \Theta(n)$. ♢

**Exercise 11.12:** We return to the previous median problem with recurrence $M(n) = M(n/5) + M(7n/10) + Cn$. In this question, we are interested in constant factors, not just asymptotics.
(a) Determine the value of $C$ in this algorithm. For this purpose, use the fact that we can find the median of five elements with 6 comparisons (Exercise in §I.3).
(b) Using Real Induction, show that $M(n) \leqslant Kn$ (ev.). Determine the optima value of $K$ as a function of $C$. ♢

*do not use our usual simplification rule to replace $C$ by 1 here!*

**Exercise 11.13:** Jack's algorithm has complexity that satisfies this recurrence:

$$Ja(n) = 2Ja(n/3) + Ja(2n/5) + n.$$

Jill's algorithm satisfies

$$Ji(n) = Ji(2n/3) + 2Ji(n/5) + n.$$

Use the Multiterm Master Theorem to decide who has the more efficient algorithm. Here is student Willa Wong's Python Script for doing these constants:

```
#!/usr/bin/python

from decimal import *
import math
def getValue(a1,b1,a2,b2):
 i = Decimal(’1’)
 while(i < 2):
  value = Decimal(a1)/Decimal(math.pow(b1,i)) \
     + Decimal(a2)/Decimal(math.pow(Decimal(b2),i)) - Decimal(’1’)
  if value < Decimal(’0.00001’):
   return i
  else:
   i += Decimal(’0.00001’)

def main():
 Tjack = getValue(Decimal(’2’), Decimal(’3’), Decimal(’1’), Decimal(’5’)/Decimal(’2’))
 Tjill = getValue(Decimal(’1’), Decimal(’3’)/Decimal(’2’), Decimal(’2’), Decimal(’5’))
 print Tjack, Tjill

if __name__ == "__main__":
    main()

```

♢

**Exercise 11.14:** Let Jack and Jill functions of the previous question be $Ja(n) = \Theta(n^\alpha)$ and $Ji(n) = \Theta(n^\beta)$. Instead of approximating $\alpha$ and $\beta$ numerically to compare them, Ravi suggests the following more geometric method of comparison (which he thinks is more insightful and avoids the use of calculators): Let

$$f(x) = 2(5^x) + 6^x, \quad g(x) = 10^x + 2(3^x)4, \quad h(x) = 15^x.$$

Then $f(\alpha) = h(\alpha)$ and $g(\beta) = h(\beta)$. It is easy to check that $\alpha, \beta$ both lies between 1 and 2.
(a) Ravi claimed that $h'(x) > g'(x) > f'(x)$, where $h'(x)$ denotes derivative with respect to $x$. Note that $h(x) = e^{x\ln(15)}$ and therefore $h'(x) = \ln(15)e^{x\ln(15)} = \ln(15)15^x$.
(b) From this we can conclude that $g(x)$ will intersect $h(x)$ at some value of $x$ that is greater than that value of at which $f(x)$ intersects $h(x)$. In other words, $\beta > \alpha$. That is, Jack's algorithm is faster than Jill's.

Your job is to make all of Ravi's arguments rigorous. Do you agree with Ravi that this is more insightful and avoid calculators? ◇

**Exercise 11.15:** Let $T(n) = 2T(n/3) + T(n/10) + 1$. Use the Master Theorem to derive a sublinear upper bound on $T(n)$. ◇

**Exercise 11.16:** Suppose $T(n) = T(n/3) + T(2n/9) + 1$. Then $T(n) = \Theta(n^\alpha)$. We want you to give the exact value of $\alpha$ (as an expression involving logs and square-roots).
HINT: recall the solution of $T(n) = T(n/4) + T(n/2) + 1$. ◇

**Exercise 11.17:** In the text, we sharpened our bounds for the conjugation tree recurrence function $T(n)$ by expanding the recurrence (125) just once, and then applying (128),
(a) Let us now expand (125) twice before applying (128). Verify that the new bounds are further improvements.
(b) Show that this improvement be repeated indefinitely? ◇

**Exercise 11.18:** Consider $T(n) = T(n/b_1) + T(n/b_2) + T(n/b_3) + 1$ where $1 < b_1 \leqslant b_2 \leqslant b_3$. What is the lower bound on $T(n)$ using (128)? Under what conditions on $b_1, b_2, b_3$ can you obtain a better bound by omitting the smallest term? ◇

End Exercises

## §12. Differencing and QuickSort

Summation is the discrete analogue of integration. Extending this analogy, we now introduce the **differencing** as the discrete analogue of differentiation. Thus differencing is the inverse of summation. The differencing operation $\nabla$ applied to any complexity function $T(n)$ yields another function $\nabla T$ defined by

$$(\nabla T)(n) = T(n) - T(n-1).$$

Differentiation often simplifies an equation: thus, $f(x) = x^2$ is simplified to the linear equation $(Df)(x) = 2x$, using the differential operator $D$. Similarly, differencing a recurrence equation for $T(n)$ may lead to a simpler recurrence for $(\nabla T)(n)$. Indeed, the "standard form" (79) can be rewritten as

$$\nabla t(n) = f(n).$$

This is just an equation involving a difference operator — the discrete analogue of a differential equation.

For example, consider the recurrence

$$T(n) = n + \sum_{i=1}^{n-1} T(i).$$

This recurrence does not immediately yield to the previous techniques. But note that

$$(\nabla T)(n) = 1 + T(n-1).$$

Hence $T(n) - T(n-1) = 1 + T(n-1)$ and $T(n) = 2T(n-1) + 1$, which can be solved by the method of range transformation.

*Solve it!*

**¶47. QuickSort.** A well-known application of differencing is the analysis of the QuickSort algorithm of Hoare. We remark that the QuickSort paradigm is extremely powerful, capable of profound generalizations to many problems in Computational Geometry. Hence it is worthwhile grasping the key ideas of this algorithm and its analysis.

In QuickSort, we randomly pick a "pivot" element $p$. If $p$ is the $i$th largest element, this subdivides the $n$ input elements into $i-1$ elements less than $p$ and $n-i$ elements greater than $p$. Then we recursively sort the subsets of size $i-1$ and $n-i$. For a detailed description of QuickSort, including a different analysis, see Chapter VIII. The recurrence is

$$T(n) \quad = \quad n + \frac{1}{n}\sum_{i=0}^{n-1}(T(i-1) + T(n-i)), \tag{133}$$

since for each $i$, the probability that the two recursive subproblems in QuickSort are of sizes $i$ and $n-i$ is $1/n$. The additive factor of "$n$" indicates the cost (up to a constant factor) to subdivide the subproblems; there is no cost in "merging" the solutions of the subproblems. The recurrence (133) is an example of a **full-history recurrence**, so-called because $T(n)$ depends on $T(m)$ for all smaller values of $m$.

Simplifying (133),

$$\begin{array}{lcll}
T(n) & = & n + \frac{2}{n}\sum_{i=0}^{n-1} T(i) & \\
nT(n) & = & n^2 + 2\sum_{i=0}^{n-1} T(i) & \text{[Multiply by } n\text{]} \\
(n-1)T(n-1) & = & (n-1)^2 + 2\sum_{i=0}^{n-2} T(i) & \text{[Substitute } n \text{ by } n-1\text{]} \\
nT(n) - (n-1)T(n-1) & = & 2n - 1 + 2T(n-1) & \text{[Differencing operator for } nT(n)\text{]} \\
nT(n) & = & 2n - 1 + (n+1)T(n-1) & \text{[Simplify]} \\
\frac{T(n)}{n+1} & = & \frac{2}{n+1} - \frac{1}{n(n+1)} + \frac{T(n-1)}{n} & \text{[Divide by } n(n+1) \text{ (range transform)]} \\
t(n) & = & \frac{2}{n+1} - \frac{1}{n(n+1)} + t(n-1) & \text{[Define } t(n) = T(n)/(n+1)\text{]} \\
 & = & 2(H_{n+1} - 1) - \sum_{i=1}^{n} \frac{1}{i(i+1)} + t(0) & \text{[Telescoping a standard form]}
\end{array}$$

Thus we see that $t(n) \leqslant 2H_{n+1}$ (assuming $t(0) = 0$) and hence

$$T(n) = 2n \ln n + \mathcal{O}(n) = \Theta(n \ln n).$$

If we are interested in the lower order term $\mathcal{O}(n)$, we can evaluate the sum $\sum_{i=1}^{n} \frac{1}{i(i+1)}$ quite sharply (see a previous Exercise).

**¶48. QuickSelect.** The following recurrence is a variant of the QuickSort recurrence, and arises in the average case analysis of the QuickSelect algorithm:

$$T(n) = n + \frac{T(1) + T(2) + \cdots + T(n-1)}{n} \tag{134}$$

In the selection problem we need to "select the $k$th largest" where $k$ is given (This problem is studied in more detail in Chapter XXX). Recursively, after splitting the input set into subsets of sizes $i-1$ and $n-i$ (as in QuickSort), we only need to continue with one of the two subsets (unlike QuickSort). This explains why, compared to (133), the only change in (134) is to replace the constant factor of 2 to 1. To solve this, let us first multiply the equation by $n$ (a range transform!). Then, on differencing, we obtain

$$\begin{aligned} nT(n) - (n-1)T(n-1) &= 2n - 1 + T(n-1) \\ nT(n) - nT(n-1) &= 2n - 1 \\ T(n) - T(n-1) &= 2 - \frac{1}{n} \\ T(n) &= 2n - \ln n + \Theta(1). \end{aligned}$$

Again, we obtain an exact solution.

**¶49. Improved QuickSort.** As QuickSort is a practical algorithm, there is interest in improving the multiplicative constants in its running time. To do this, we first randomly choosing three elements, and picking the median of these three to be our pivot. The resulting recurrence is slightly more involved:

$$T(n) = n + \sum_{i=2}^{n-1} p_i[T(i-1) + T(n-i)] \tag{135}$$

where

$$p_i = \frac{(i-1)(n-i)}{\binom{n}{3}}$$

is the probability that the pivot element gives rise to subproblems of sizes $i-1$ and $n-i$. See Chapter 8 on Probabilistic Analysis where we further discuss QuickSort.

## Exercises

**Exercise 12.1:** Consider the following recurrence:

$$U(n) = n + \max_{m=1}^{n} \{U(m-1) + U(n-m)\}$$

We want to show that under DIC, $U(n) = \Omega(n^2)$. Here are some pitfalls that students encounter:

(a) Student A says: assume by DIC that $U(m) \geqslant m^2$. Then $U(n) \geqslant n + U(0) + U(n-1) \geqslant n + 0^2 + (n-1)^2 = n + (n-1)^2$. Clearly, $n + (n-1)^2 = \Omega(n^2)$, QED.

(b) Student B says: assume by DIC that $U(m) \geqslant Cm^2$ for some $C > 0$. Then $U(n) = n + U(m-1) + U(n-m) \geqslant n + C(m-1)^2 + C(n-m)^2$ where the first equality follows by choosing the optimum value of $m$. Then a sequence of *algebraic manipulations only*, the student concludes that $U(n) \geqslant Cn^2$.

(c) Student C considers the following function of $m$: $f(m) := n + C(m-1)^2 + C(n-m)^2$ (fixing $n$). If we assume by DIC that $U(m) \geqslant Cn^2$, it follows that $U(n) \geqslant \max_{m=1}^{n} f(m)$. Student C noted that $f(m)$ is unimodal with a minima at $m = (n-1)/2$, and concluded that $U(n) \geqslant f((n-1)/2)$.

Please provide appropriate advice to Students A, B and C. ♢

**Exercise 12.2:** Solve the following recurrences to Θ-order:

$$T(n) = n + \frac{2}{n} \sum_{i=\lfloor n/2 \rfloor}^{n-1} T(i).$$

HINT: Because of the upper bound $\lfloor n/2 \rfloor$, the function $\nabla T(n)$ has different behavior depending on whether $n$ is even or odd. Simple differencing does not seem to work well here. Instead, we suggest the guess and verify-by-induction approach. ♢

**Exercise 12.3:** Generalize the previous question. Consider the recurrence

$$T(n) = n + \frac{c}{n} \sum_{i=1+\lfloor \alpha n \rfloor}^{n-1} T(i)$$

where $c > 0$ and $0 \leqslant \alpha < 1$ are constants.
(a) Solve the recurrence for $c = 2$.
(b) Solve $T(n)$ when $c = 4$ and $\alpha = 0$.
(c) Fix $c = 4$. Determine the range of $\alpha$ such that $T(n) = \Theta(n)$. You need to argue why $T(n)$ is not $\Theta(n)$ for $\alpha$ outside this range.
(d) Determine the solution of this recurrence for general $c, \alpha$. ◇

**Exercise 12.4:** (a) Suppose that in the base case of QuickSort, we do nothing whenever the size of the subarray to be sorted has 10 or less keys. Call this "`QuirkSort`".
(i) Describe the nature of the output from `QuirkSort`.
(ii) Describe a linear time method to take the output of `QuirkSort` and make it into a sorted array.
(iii) Explain why your method in (ii) takes linear time. ◇

**Exercise 12.5:**
(a) Show that every polynomial $p(X)$ of degree $d$ can be written as a sum of binomial coefficients with suitable coefficients $c_i$:

$$p(X) = c_d \binom{X}{d} + c_{d-1} \binom{X}{d-1} + \cdots + c_1 \binom{X}{1} + c_0.$$

(b) Assume the above form for $p(X)$, express $(\nabla p)(X)$ as a sum of binomial coefficients. HINT: what is $\nabla \binom{m}{n}$? ◇

End Exercises

## §13. Other Recurrences

There is a wide variety of recurrences which we have barely hinted at. For instance, the typical recurrences arising in counting combinatorial structures have an exponential (e.g., $T(n) = 2T(n-1) + f(n)$) or double exponential growth (e.g., $T(n) = T(n-1)^2 + f(n)$). See Knuth for such examples. In this section, we focus on some other types of recurrences.

### §13.1. Recurrences with Max or Min

Many recurrences in computer science involve the Max or Min operation. Here we give three examples.

**¶50. QuickSort Variant.** Consider the following variant of QuickSort: each time after we partition the problem into two subproblems, we will solve the subproblem that has the smaller size first (if their sizes are equal, it does not matter which order is used). We want to analyze the depth of the recursion stack. If a problem of size $n$ is split into two subproblems of sizes $n_1, n_2$ then $n_1 + n_2 = n - 1$. Without loss of generality, let $n_1 \leqslant n_2$. So $0 \leqslant n_1 \leqslant \lfloor (n-1)/2 \rfloor$. If the stack contains problems of sizes $(n_1 \geqslant n_2 \geqslant \cdots \geqslant n_k \geqslant 1)$ where $n_k$ is the problem size at the top of the stack, then we have

$$n_{i-1} \geqslant n_i + n_{i+1}.$$

Since $n_1 \leqslant n$, this easily implies $n_{2i+1} \leqslant n/2^i$ or $k \leqslant 2 \lg n$. A tighter bound is $k \leqslant \log_\phi n$ where $\phi = 1.618\ldots$ is the golden ratio. This is not tight either.

The depth of recursion satisfies

$$D(n) = \max_{n_1=0}^{\lfloor (n-1)/2 \rfloor} [\max\{1 + D(n_1), D(n_2)\}]$$

This recurrence involving max is actually easy to solve. Assuming $D(n) \leqslant D(m)$ for all $n \leqslant m$, and for any real $x$, $D(x) = D(\lfloor x \rfloor)$, it is easy to see that $D(n) = 1 + D(n/2)$. Using the fact that $D(1) = 0$, we obtain $D(n) \leqslant \lg n$. [Note: $D(1) = 0$ means that all problems on the stack has size $\geqslant 2$.

**¶51. Solving a Problems on a Binary Tree.** Consider this recurrence which involves both Max and Min:

$$C(n) = \max_{m \geqslant 0}^{n-1} \{C(m) + C(n - m - 1) + \min\{m, n - m - 1\}\} \tag{136}$$

This represent the cost to solve a recursive problem represented by a binary tree $T$ on $n$ nodes, where the left and right subtrees have sizes $m$ and $n - m - 1$, respectively. To solve the problem on $T$, we recursively solving the problem on the left and right subtrees, and then marry the two sub-solutions at a cost of $\min\{m + 1, n - m\}$. We view (136) as a real recurrence. In the maximization notation of this recurrence, the variable $m$ as ranges[21] over all real values between 0 and $n - 1$.

Recall the $n$th Harmonic number for any real $n$ is defined as $H_n = \sum_{i \geqslant 1}^{n} i^{-1}$ (using the descending sum convention).

**Lemma 16** *Assuming DIC,*

$$C(n) \leqslant nH_n(\text{ev.}). \tag{137}$$

*Proof.* By DIC, we can assume (137) is true for all $n \leqslant n_0$ $(n_0 \geqslant 1)$. Then we have

$$\begin{array}{rcll}
C(n) & = & C(m) + C(n - m - 1) + \min\{m, n - m - 1\} & \text{(for optimal } m\text{)} \\
& \leqslant & mH_m + (n - m - 1)H_{n-m-1} + \min\{m, n - m - 1\} & \text{(by induction hypothesis)} \\
& \leqslant & (n - 1)H_m + (n - m - 1) & \text{(WLOG, } m \geqslant n - m - 1\text{)} \\
& \leqslant & (n - 1)\sum_{i \geqslant 1}^{m} i^{-1} + (n - m - 1) & \text{(using the descending sum convention)} \\
& \leqslant & (n - 1)\sum_{i \geqslant 1}^{m} i^{-1} + \sum_{i \geqslant m}^{n} i^{-1}(n - 1) & \text{(each term in the latter sum is} \geqslant 1\text{)} \\
& \leqslant & (n - 1)\sum_{i \geqslant 1}^{n} i^{-1} & \text{(the first } \lfloor m \rfloor \text{ terms may have increased)} \\
& = & (n - 1)H_n. &
\end{array}$$

**Q.E.D.**

This proves $C(n) = O(n \log n)$. This bound exploits the Min in (136). For instance, if we replace the Min by a Max, then the solution is $C(n) = \Theta(n^2)$ (Exercise). This $O(n \log n)$ bound is instructive: in effect, it says that the worst case value of $m$ in (136) is when $m \sim n/2$, thus reducing the recurrence to look like $C(n) = 2C(n/2) + n$, yielding the $\Theta(n \log n)$ solution. So the Min has the effect of ensuring that the balanced binary tree $T$ is the worst case solution.

Fredman [6] considered the general class of recurrences of the form

$$M(n) = g(n) + \min_{0 \leqslant k \leqslant n-1} \{\alpha M(k) + \beta M(n - k - 1)\}$$

which arises from analysis of binary search trees.

[21] Alternately, under the descending sum convention, we would interpret $m$ as ranging over the discrete set of values $n - 1, n - 2, \ldots, \{n\}$ where $\{n\}$ is the fractional part of $n - 1$. But by allowing $m$ to vary over all reals from 0 to $n - 1$, we obtain a stronger result in our upper bound for $C(n)$.

**¶52. Analysis of $\epsilon$-Nets.** The following recurrence arise in the analysis of a class of data structures called $\epsilon$-nets, first studied by Haussler and Welzl. Fix $0 < \epsilon < 1$ and $m \geqslant 2$. By a **partition** of a real number $n > 1$, we mean a multiset $P$ of real numbers such that $n \leqslant \sum_{v \in P} v$ and each $1 \leqslant v \leqslant 1$. The size $|P|$ of the partition is the number elements in $P$.

$$T(n) = 1 + \max_P \sum_{v \in P} T(v) \tag{138}$$

where $P$ ranges over all partitions of $\epsilon n$ of size $\leqslant m$. There is a trivial solution to this: the constant function

$$T(n) = 1/(1-m)$$

for all $n$. But $T(n) < 0$ in this case and we seek a non-negative solution. Assuming that $T(n)$ is a convex cap[22], it is easy to see that

$$T(n) = 1 + mT(\epsilon n/m).$$

By the Master Theorem, the watershed constant is $w = \log_b a = \log_{m/\epsilon} m < 1$, and the recurrence has solution

$$T(n) = \Theta(n^{\log_{m/\epsilon} m}).$$

To show $T(n)$ is a convex cap, we note that it is continuous (Exercise) and a monotonic non-decreasing function. Then it suffices (Exercise) to prove that

$$T(x) + T(y) \leqslant 2T((x+y)/2) \tag{139}$$

where we now regard $T(x)$ as a real function defined for all $x \geqslant 0$. This turns out to be easy to show inductively, assuming the base case where $T(x) = x$ (or $T(x) = 0$) for all $0 \leqslant x \leqslant 1$.

## §13.2. A Log-square Solution

Consider the recurrence

$$T(n) = 1 + T(n - \frac{n}{\log n}). \tag{140}$$

This does not yield to our standard techniques. To probe deeper, note some simple bounds. It is easy to see that $T(n) \leqslant n$ since this is the solution to the recurrence $T(n) \leqslant 1 + T(n-1)$. Likewise $T(n) \geqslant \lg n$ since this is the solution to $T(n) \geqslant 1 + T(n/2)$.

To get a better upper bound, we note that

$$\begin{aligned} T(n) &= 1 + T\left(n\left(1 - \frac{1}{\log n}\right)\right) \\ &\leqslant 2 + T\left(n\left(1 - \frac{1}{\log n}\right)^2\right), \qquad (\mathit{why?}) \\ &\vdots \\ &\leqslant k + T\left(n\left(1 - \frac{1}{\log n}\right)^k\right) \end{aligned}$$

using monotonicity of $T(n)$. Hence $T(n) = k$ if we assume $T(n) = 0$ for $n \leqslant 1$ and $k$ is chosen so that

$$\left(1 - \frac{1}{\log n}\right)^k \leqslant 1/n < \left(1 - \frac{1}{\log n}\right)^{k+1}.$$

Taking natural logs, and assuming for simplicity that $\log = \ln$ in (140), we see that

$$\begin{aligned} (k+1)\ln\left(1 - \frac{1}{\ln n}\right) &> -\ln n, \\ (k+1)\left(-\frac{1}{\ln n}\right) &> -\ln n, \qquad (\text{since} \quad \ln(1+x) \leqslant x \text{ for } \ |x| < 1), \\ k+1 &< \ln^2 n. \end{aligned}$$

[22] We say a real function $f(x)$ is **convex cap** if for all $0 < \alpha < 1$, $f(x) + f(y) \leqslant 2f(\alpha x + (1-\alpha)y)$. For completeness, we say $f(x)$ is **convex cup** if for all $0 < \alpha < 1$, $f(x) + f(y) \geqslant 2f(\alpha x + (1-\alpha)y)$.

Up to a constant factor, this is also the lower bound: we show that $T(n) \geqslant C\ln^2 n$ by induction:

$$\begin{aligned}
T(n) &\geqslant 1 + C\ln^2\left(n\left(1 - \frac{1}{\log n}\right)\right) \\
&= 1 + C(\ln n + \ln\left(1 - \frac{1}{\log n}\right))^2 \\
&\geqslant 1 + C(\ln n - \frac{2}{\ln n})^2, \qquad \text{since } \ln(1+x) \geqslant x - x^2/2 \text{ for } |x| < 1 \\
&\geqslant C\ln^2 n.
\end{aligned}$$

Thus $T(n) = \Theta(\ln^2 n)$.

REMARK: If we were told from the beginning to verify that $T(n) = \Theta(\ln^2 n)$, this would be routine. What we are demonstrating here is the process of discovering that $\Theta(\ln^2 n)$ is the correct answer.

## Exercises

**Exercise 13.1:** The following recurrence $S(n)$ arises in many situations, for instance, in analyzing the space complexity of partition trees. Recall that a **partition** of $n$ is any multiset $I$ of real numbers such that $n \leqslant \sum_{v\in I} v$ where $v \geqslant 1$ for each $v \in I$. E.g., $n = 4.7$ and $I = \{1, 1, 2.7\}$. Consider the recurrence $S(1) = 1$ and for $n > 1$,

$$S(n) = 1 + \max_I \sum_{v\in I} S(v)$$

where $I$ ranges over all partitions of $n$. Prove that $S(n) = O(n)$. $\Diamond$

**Exercise 13.2:** The **4-Tower of Hanoi Problem** is this: we have $n$ discs, no two discs of the same size. There are 4 spots on the ground where these discs may be stacked: $A, B, C, D$. The set of discs at any spot must form a **pile** (i.e., the largest disc at the bottom of the pile, followed by next largest, and so on, until the smallest disc at the top). Initially, all $n$ discs are in a pile $A$ (so piles $B, C, D$ are empty). GOAL: *to move all the discs in pile $A$ to pile $B$.* To achieve this GOAL, we are allowed to move the top disc in one pile to the top of any destination pile (the moved disc must be smaller than the top disc of the destination pile). Let $T_4(n)$ denote the minimum number of such moves necessary to achieve GOAL.
(a) As warm up, consider the original Tower of Hanoi problem: here, we are allowed only three piles $A, B, C$. The optimum number of moves in this case is $T_3(n) = 2^n - 1$. Prove this (we need upper and lower bounds).
(b) Prove an upper bound of $T_4(n) = 2^{O(\sqrt{n})}$. HINT: You may use the result of Part(a), and also note that $2^{\sqrt{n}} \times 2^{\sqrt{n}} = 2^{O(\sqrt{n})}$. $\Diamond$

**Exercise 13.3:** Solve for $C(n)$ where

$$C(n) = \max_{m=0,\ldots,n-1} \{C(m) + C(n-m-1) + \max\{m+1, n-m\}\}.$$

Note that this is similar to (136) except that the Min has been replaced by a Max. $\Diamond$

**Exercise 13.4:** Try to obtain tight constants for the recurrence (140). What if log is not the natural logarithm in the original equation? $\Diamond$

**Exercise 13.5:** Show that $T(x)$ in (139) is continuous by exploiting the fact that the addition and maximum functions are continuous. $\Diamond$

**Exercise 13.6:** Prove that if $T(x)$ is continuous and satisfies equation (139) then it is a convex cap. ◇

**Exercise 13.7:** Bound the solution to the recurrence $T(n) = T(n-1) + 2T(n/2) + n$. This is an interesting mixture of linear recurrence and the master recurrence. ◇

**Exercise 13.8:** (Leighton 1996) Show that $T(n) = 2T(\frac{n}{2} - \frac{n}{\lg n})$ has solution $T(n) = \Theta(n \log^{\Theta(1)} n)$. Assume that $T(n) = 1$ for $n \leqslant 5$, and the recurrence holds for $n > 5$. Thus $T(5 + \varepsilon) = 2$, so this function is discontinuous. ◇

**Exercise 13.9:** Analyze the behavior of the function $T(n)$ defined by the recurrence $T(n) = nT(\log n)$. Give upper and lower bounds for $T(n)$ using "closed form expressions" in terms of the functions $\log^{(i)} n$, $i \geqslant 0$. **Note:** This recurrence arises in an early version of the fast integer multiplication algorithm of Schönhage and Strassen. ◇

**Exercise 13.10:** Solve the recurrence $T(n) = 1 + \max_{(n_1,n_2,n_3,n_4)}\{T(n_1) + T(n_2) + T(n_3) + T(n_4)\}$ where $(n_1, \ldots, n_4)$ ranges over all non-negative numbers such that $\sum_{i=1}^4 n_i = \frac{3n}{2}$ and each $n_i \leqslant n/2$. ◇

**Exercise 13.11:** Solve the following recurrences to $\Theta$-order:
(a) $T(n) = 1 + 2T(n - \frac{n}{\log n})$.
(b) $T(n) = 2^n T(n/2) + n^n$.
(c) $T(n) = 1 + T(\frac{n}{\log n})$.
HINT: these recurrences are considerably harder than most of what we encounter. First guess non-tight upper and lower bounds and verify by induction. Then try to tighten these bounds. ◇

End Exercises

## §13.3. Multivariable Recurrences

So far, our recurrences involve only one variable. But multivariable recurrences arise in several ways: one source of such recurrences is multidimensional problems in computational geometry (one of the variable is the dimension).

The pre-processing problem of **point dominance queries** in $d$-dimensions is as follows: given a set $S \subseteq \mathbb{R}^d$ of $n$ points, construct a data structure $D(S)$ such that for any query point $p \in \mathbb{R}^d$, we can quickly determine if there is any point $x \in S$ that **dominates** $p$ (this means $x \geqslant p$, componentwise). One solution is to pick some $c \in R$ such that $S$ splits into two subsets $S_1, S_2$ of size $n/2$ each, where the first component of each $x \in S_1$ is $\leqslant c$, and the first component of each $x' \in S_2$ is $\geqslant c$. To answer the query for $p$, begin by comparing the first component $p_1$ of $p$ to $c$: if $p_1 > c$ then it is sufficient to recursively check if some $x \in S_2$ dominates $p$. If $p_1 \leqslant c$, we must do two searches: (i) check if some $x \in S_1$ dominates $p$ and (ii) check if some $x \in S_2$ dominates $p$. The search in (i) is, however, done in $d-1$ dimensions since we may ignore the first components. Thus the time for answering queries satisfies the recurrence

$$T(n,d) = 1 + T(n/2,d) + T(n/2,d-1).$$

It is not hard to see that $T(n,1) = \mathcal{O}(1)$. Then we may verify the solution $T(n,d) = \Theta(\log^{d-1} n)$.

**¶* 53. Output-sensitive algorithms.** ultivariable recurrences arise in the analysis of "output-sensitive" algorithms. Such algorithms has, besides the traditional **input parameter** $n$, an (implicit) **output parameter** $h$, which is the measures the size of the output for the given input instance. The computational complexity of such algorithms depends on both $n$ and $h$. An example is the problem of computing the convex hull of a set of $n$ points in the plane. The output size is just the number of points in the actual convex hull. There are well-known $\mathcal{O}(n \log n)$ algorithms for this problem. Kirkpatrick and Seidel has given an algorithm whose time complexity satisfies the following recurrence:

$$T(n,h) = \mathcal{O}(n) + \max_{h_1+h_2=h-1} \left\{ T(\frac{n}{2}, h_1) + T(\frac{n}{2}, h_2) \right\}.$$

Here, $h_i$ are positive integers. We may assume $T(n,h) = \mathcal{O}(n)$ for $h \leqslant 3$. To see that $T(n,h) = \mathcal{O}(n \log h)$, we could of course just substitute and verify. But it is more instructive to argue as follows: consider a "recursion tree" corresponding to a possible expansion of the recurrence relation for $T(n,h)$. There are exactly $h$ nodes in this binary tree, where each internal node at depth $i$ (the root is depth 0) carries a "cost" of $n/2^i$. The "cost" of the tree just the sum of these costs at the internal nodes. So $T(n,h)$ is the maximum cost over all possible recursion trees. The *claim* $T(n,h) = \mathcal{O}(n \log h)$ follows if we prove that the maximum cost occurs when the tree has depth at most $\log_2 h$ (since the total cost of all nodes at any depth $i$ is invariably $n$). For the sake of contradiction, suppose we have a maximum cost tree with depth $d > \log_2 h$. Then there is a node at depth $d-1$ whose children are leaves at depth $d$. We can transfer these two children to become the children of some other node at depth $\leqslant d-2$. This would increase the cost for the tree, contradiction.

Exercises

**Exercise 13.12:** Show that if $S(n,d)$ is the space requirement for the above data structure, then $S(n,d) = 1 + 2S(n/2,d) + S(n/2,d-1)$. Solve this recurrence. What is $S(n,1)$? ◇

**Exercise 13.13:** Consider the following recurrence

$$T(n,h) = \mathcal{O}(n) + \max_{h_1+h_2=h-1; c_1+c_2=1} \{T(c_1 n, h_1) + T(c_2 n, h_2)\}.$$

(a) Solve for $T(n,h)$ with only the assumption $h_i \geqslant 1, c_i > 0$ in the above.
(b) Solve for $T(n,h)$ with the *additional* assumption that $c_i \leqslant \alpha$ where $0 < \alpha < 1$ is fixed. Generalize the above argument about the shape of the maximum cost recursion tree. ◇

**Exercise 13.14:** (Sharir-Welzl) The following recurrence arises in analyzing the diameter of $n$-dimensional polytopes with $m$ facets:

$$f(n,m) = f(n-1,m-1) + \frac{2}{m} \sum_{i=1}^{m} f(n-1,i).$$

Solve the recurrence. ◇

**Exercise 13.15:** (Simultaneous Recurrences) Consider the following mutual recurrences from Korenblit and Levit (2012) involving three complexity functions $T_0, T_1, T_3$:

$$\begin{array}{rcl} T_0(n) & = & 1 + 2T_0(n/2) + 2T_1(n/2) + 2T_2(n/2) \\ T_1(n) & = & 1 + T_0(n/2) + 3T_1(n/2) + 2T_2(n/2) \\ T_2(n) & = & 1 + 2T_1(n/2) + 4T_2(n/2). \end{array}$$

These recurrences arise in a problem to compute the algebraic expression associated with a family of directed graphs. Solve up to $\Theta$-order. ◇

End Exercises

## §14. Orders of Growth

This section is a practical one, giving you some tools to determine domination-type relation between pairs of functions.

Jack: *My algorithm has time complexity* $\Theta((\lg n)^n)$.

Jill: *Ah, but mine runs in* $\Theta(n^{\lg n})$.

Who has the better (i.e., faster) algorithm – Jack or Jill? Most students would not be able to tell the answer right away. Let $T_K(n) = (\log n)^n$ and $T_L(n) = n^{\lg n}$ be the complexity of Jac$\underline{\text{k}}$'s and Jil$\underline{\text{l}}$'s algorithms. Instead of comparing $T_K(n)$ and $T_L(n)$ we can compare their logarithms:

$$\lg T_K(n) = n \lg\lg n \quad \text{versus} \quad \lg T_J(n) = \lg^2 n. \tag{141}$$

If you still do not see any domination relation, you can take logs again:

$$\lg n + \lg\lg\lg n \quad \text{versus} \quad 2\lg\lg n. \tag{142}$$

It is now clear that the left-hand side super-dominates the right-hand side, since

$$\lg n \ggg \lg\lg n. \tag{143}$$

Working backwards to the original comparison, we conclude that

$$T_K(n) = (\lg n)^n \ggg T_L(n) = n^{\lg n}. \tag{144}$$

Thus Jack's algorithm is slower than Jill's.

*Be careful! "faster growing complexity"* $\equiv$ *"slower algorithm"*

**¶54. On the Heuristic of taking Logs.** Is the above argument rigorous? The heuristic of taking logs amounts to an application of the following "reverse" inference rule:

$$(f \ggg g) \Longleftarrow (\lg f \ggg \lg g). \tag{145}$$

Here, "$A \Longleftarrow B$" reads "$A$ holds *provided* $B$ holds". Logically, $A \Longleftarrow B$ and $B \Longrightarrow A$ are equivalent, but the backwards formulation seems more natural in the proofs of domination-type relations.

**The two modes of proof**: Given propositions $A$ and $B$, if "$A \Longleftarrow B$", we say "$A$ *provided* $B$". And if "$B \Longrightarrow A$", we say "$B$ *implies* $A$". Although these two assertions about $A$ and $B$ are logically equivalent, they give rise to two very distinct modes of organizing a proof: $A_1 \Longleftarrow A_2 \Longleftarrow A_3 \Longleftarrow \cdots \Longleftarrow A_n$ ("provisional mode") versues $A_n \Longrightarrow A_{n-1} \Longrightarrow A_{n-2} \Longrightarrow \cdots \Longrightarrow A_1$ ("implicational mode"). In a long chain ($n$ is large), the provisional mode seems better since you begin with what you really want to prove ($A_1$) and each proposition $A_i$ is introduced as needed. In the implicational mode, you begin with $A_n$ which may appear to be out of the blue if your target is $A_1$. In a complete proof, the provisional chain ends with $A_n = \texttt{true}$ or some axiom or a given. Similarly, the implicational chain begins with $A_n = \texttt{true}$, etc. More generally, these chains are replaced by proof trees or DAGs. Interested students may look up the subject of proof theory in mathematical logic or the theorem proving literature. See the logic notes in Chapter I (Appendix A).

The use of $\log T(n)$ as a surrogate for $T(n)$ for comparisons seems to be justified because log is a monotone increasing function: $x \leqslant y$ iff $\log x \leqslant \log y$. You might want to review the properties of exponentials and logarithms in Appendix A. More precisely, if $a$ and $b$ are real, then

$$a \geqslant b \iff \log a \geqslant \log b$$

provided $b > 0$. If “$\succcurlyeq$” is one of $\succeq, \succ, \ggg$, is there a similar equivalence? Instead of equivalence, we only want the reverse inference,

$$(f \text{ “}\succcurlyeq\text{” } g) \Longleftarrow (\log f \text{ “}\succcurlyeq\text{” } \log g).$$

We will focus on the case where “$\succcurlyeq$” is $\ggg$, leaving $\succeq$ and $\succ$ to the Exercises. Unfortunately[23] when “$\succcurlyeq$” is $\ggg$, the rule (145) is not sound! Here is a counter example: let $g = 1$ and $f = 2$. Then $1 = \lg f \ggg \lg g = 0$, but it is not true that $f \ggg g$. The following lemma overcomes this by an additional guarantee that $\lg f$ is growing fast enough:

**Lemma 17** *Let $f, g$ be complexity functions.*
*If $\lg f$ super-dominates both 1 and $\lg g$, then $f$ super-dominates $g$. In symbols,*

$$(f \ggg g) \Longleftarrow (\lg f \ggg \lg g) \wedge (\lg f \ggg 1). \tag{146}$$

*Proof.*

$$\begin{array}{llll}
(f \ggg g) & \iff & (\forall C > 0)[Cf > g \text{ (ev.)}] & \\
 & \iff & (\forall C \in (0,1))[Cf > g \text{ (ev.)}] & \text{[Can restrict } C \text{ to } 0 < C < 1] \\
 & \iff & (\forall C \in (0,1))[\lg C + \lg f > \lg g \text{ (ev.)}] & \text{[Taking logs of both side]} \\
 & \iff & (\forall C \in (0,1))[(C \lg f \geqslant \lg g) \wedge (\lg C + (1-C)\lg f > 0) \text{ (ev.)}] & \text{[Split } \lg f \text{ into two parts]} \\
 & \Longleftarrow & (\forall C \in (0,1))[(C \lg f \geqslant \lg g \wedge \lg f > \left(\frac{-\lg C}{1-C}\right) \text{ (ev.)})] & \text{[rearranging]} \\
 & \Longleftarrow & (\lg f \ggg \lg g) \wedge (\lg f \ggg 1) & \text{[(suffices that } \lg f \text{ is unbounded)]}
\end{array}$$

**Q.E.D.**

We view, “$\lg f \ggg 1$” as the subsidiary clause that is sufficient to give us a valid reverse inference. Returning to our heuristic Jack and Jill argument, we see that (145) can be fully justified provided we can justify the subsidiary clauses: “$n \lg\lg n \ggg 1$” for (141), and “$\lg n \ggg 1$” for (142). This follows from a general fact:

**Lemma 18** *For all $k \in \mathbb{N}$, we have*

$$\lg^{(k)}(n) \ggg \lg^{(k+1)}(n) \ggg 1.$$

*where*

$$\lg^{(k)}(n) = \begin{cases} n & \text{if } \ k = 0, \\ \lg(\lg^{(k-1)}(n)) & \text{if } \ k > 0. \end{cases}$$

We could generalize this lemma to $k \in \mathbb{Z}$ by defining

$$\lg^{(k)}(n) = 2^{\lg^{(k+1)}(n)}$$

when $k < 0$. Also, we can extend to logarithms to any fixed base $b > 1$.

[23] Thanks to Xu Cao (xc2057@nyu.edu) (fall 2020, NYU Shanghai) for pointing out the bug.

**¶55. Some rules for comparing functions.** When functions falls outside these well-known categories, we can use some rules to help us compare them. Here are two simple rules for comparing functions up to $\Theta$-order:

**(SR) Sum Rule:** In a "direct" comparison involving a sum $f(n) + g(n)$, ignore the smaller term in this sum.
E.g., in comparing $n^2 + n \log n + 5$, you should ignore the "$n \log n + 5$" term.

**(PR) Product Rule:** If $0 \preceq f \preceq f'$ and $0 \preceq g \preceq g'$ then $fg \preceq f'g'$.
E.g., this rule implies $n^b \prec n^c$ when $b < c$ (since $1 \prec n^{c-b}$, by the logarithm rule next).

Another way to compare functions is to compare their exponents instead:

**(LR) Log Rule:** $1 \ll \log^{(k+1)} n \ll \log^{(k)} n$ for any integer $k \geqslant 0$. Here $\log^{(k)} n$ refers to the $k$-fold application of the logarithm function and $\log^{(0)} n = n$.

**(ER) Exponentiation Rule:** We have two versions: assume $0 \leqslant f$.
(ER1) If $f \preceq g$ then $2^f \preceq 2^g$.
(ER2) If $f \ll g$ then $2^f \ll 2^g$.
The constant 2 can be replaced by any $d > 1$.

**Example.** Suppose we want to compare $n^{\log n}$ versus $(\log n)^n$. According to the Exponentiation Rule (ER), $n^{\log n} \prec (\log n)^n$ follows if we take logs and show that $1 \leqslant \log^2 n \leqslant 0.5 n \log\log n$ (ev.) (i.e., choose $c = 0.5$ in (ER)). In fact, we show the stronger $\log^2 n \ll n \log\log n$. Taking logs again, and by the rule of sum, it is sufficient to show $2 \log\log n \prec \log n$. Taking logs again, and by the rule of sum again, it is suffices to show $\log^{(3)} n \prec \log^{(2)} n$. But the latter follows from the rule of logarithms.

**¶* 56. On L-functions.** The functions $(\lg n)^n$ and $n^{\lg n}$ of Jack and Jill are examples of **logarithmico-exponential functions** (*L-functions* for short) from G.H. Hardy [9]. An $L$-function $f(x)$ is real and defined for all $x \geqslant x_0$ for some $x_0$ depending on $f$. They are inductively defined as either the identity function $x$, or a constant $c \in \mathbb{R}$, or else obtained as a finite composition of the functions

$$A(x), \qquad \ln(x), \qquad e^x$$

where $A(x)$ denotes[24] a real branch of an algebraic function. For instance, $A(x) = \sqrt{x}$ is the function that picks the real square-root of $x$. But we could also have taken the negative branch of the square-root.

*good time to review exponentials and logarithms in the Appendix!*

We say a set of functions is **totally ordered** if, for any $f, g$ in the set, either $f \preceq g$ or $g \preceq f$. A theorem[25] of Hardy [9] says that the set of $L$-functions is totally ordered: *if $f$ and $g$ are $L$-functions then $f \leqslant g$ (ev.) or $g \leqslant f$ (ev.).* In particular, each $L$-function $f$ is eventually non-negative, $0 \leqslant f$ (ev.), or non-positive, $f \leqslant 0$ (ev.). This is a very nice property of $L$-functions. Unfortunately, many common functions that are not $L$-functions. For instance, the sine function is not an $L$-function because neither $\sin x \geqslant 0$ (ev.) nor $\sin x \leqslant 0$ (ev.) holds. Here are some categories of $L$-functions you often encounter:

| CATEGORY | SYMBOL | EXAMPLES |
|---|---|---|
| vanishing term | $o(1)$ | $\frac{1}{n}$, $2^{-n}$ |
| constants | $\Theta(1)$ | $1$, $2 - \frac{1}{n}$ |
| polylogs | $\log^k n$ (for any $k > 0$) | $H_n$, $\log^2 n$ |
| polynomials | $n^k$ (for any $k > 0$) | $n^3$, $\sqrt{n}$ |
| super-polynomials | $n^{\Omega(1)}$ | $n!$, $2^n$, $n^{\log\log n}$ |

[24] An algebraic function $A(x)$ satisfies a polynomial equation $P(x, A(x)) = 0$ where $P(x, y)$ is a bivariate polynomial with integer coefficients. For instance $A(x) = \sqrt{x}$ is algebraic since $P(x, A(x)) = 0$ for the polynomial $P(x, y) = x^2 - y$.

[25] In the literature on $L$-functions, the notation "$f \preceq g$" actually means $f \leqslant g$ (ev.). There is a deep theory involving such functions, with connection to Nevanlinna theory.

Actually, $n!$ and $H_n$ are not $L$-functions, but they can be closely approximated by $L$-functions. The last category forms a grab-bag of anything growing faster than polynomials. These 5 categories form a hierarchy of strictly increasing $\Theta$-order.

## Exercises

**Exercise 14.1:** Consider the expression $E(n) := f(n)^{g(h(n))}$ where $\{f, g, h\} = \{2^n, 1/n, \lg n\}$. These are $6 = 3!$ possibilities for $E(n)$:

| $E(n)$ | $f$ | $g$ | $h$ |
|---|---|---|---|
| $E_1$ | $2^n$ | $1/n$ | $\lg n$ |
| $E_2$ | $2^n$ | $\lg n$ | $1/n$ |
| $E_3$ | $\lg n$ | $2^n$ | $1/n$ |
| $E_4$ | $\lg n$ | $1/n$ | $2^n$ |
| $E_5$ | $1/n$ | $2^n$ | $\lg n$ |
| $E_6$ | $1/n$ | $\lg n$ | $2^n$ |

Determine the domination relation between these functions. ♢

**Exercise 14.2:** (i) Simplify the following expressions:

$$\text{(a) } n^{1/\lg n}, \qquad \text{(b) } 2^{2^{(\lg\lg n)-1}}, \qquad \text{(c) } \sum_{i=0}^{k-1} 2^i,$$
$$\text{(d) } 2^{(\lg n)^2}, \qquad \text{(e) } 4^{\lg n}, \qquad \text{(f) } (\sqrt{2})^{\lg n}.$$

Be sure to show your simplification steps.

(ii) Re-do the above, replacing each explicit or implicit occurrence of "2" in the previous expressions by a constant $c > 1$. We view lg as $\log_2$, 4 as $2^2$ and $\sqrt{n}$ as $n^{1/2}$. Thus these expressions becomes $\log_c$, $c^c$ and $n^{1/c}$.

♢

**Exercise 14.3:** Order these in increasing big-Oh order:

(a) $n \lg n$, (b) $n^{-1}$, (c) $\lg n$, (d) $n^{\lg n}$, (e) $10n + n^{3/2}$, (f) $\pi^n$, (g) $2^n$, (h) $2^{\lg n}$.

We want you to briefly justify your ordering: use relations such as $1 \prec (\lg n)^a \prec n^b$ for all $a, b > 0$. If the answer is (a¡b¡c), you need to justify why $a < b$ and $b < c$.
Note on Grading: there are $\binom{8}{2} = 28$ pairs of functions; we grade by counting the number correct pairs in your answer. E.g., if the answer is (a,b,c) and you write (c,a,b), you have only one correct pair. But (a,c,b) has 2 correct pairs. ♢

**Exercise 14.4:** Order the following 5 functions in order of increasing $\Theta$-order:
(a) $\log^2 n$, (b) $n/\log^4 n$, (c) $\sqrt{n}$, (d) $n2^{-n}$, (e) $\log\log n$. ♢

**Exercise 14.5:** Order the following functions (be sure to parse these nested exponentiations correctly):
(a) $n^{(\lg n)^{\lg n}}$, (b) $(\lg n)^{n^{\lg n}}$, (c) $(\lg n)^{(\lg n)^n}$, (d) $(n/\lg n)^{n^{n/(\lg n)}}$. (e) $n^{n^{(\lg n)/n}}$. ♢

**Exercise 14.6:** Order the following set of 36 functions in non-increasing order of growth. Between consecutive pairs of functions, insert the appropriate ordering relationship: $\preceq$, $\asymp$, $\leqslant$ (ev.), $=$.

| | a | b | c | d | e | f |
|---|---|---|---|---|---|---|
| 1. | $\lg\lg n$ | $(\lg n)^{\lg n}$ | $2^n$ | $2^{\lg n}$ | $2^{\lg^* n}$ | $2^{2^{n+1}}$ |
| 2. | $(1/3)^n$ | $n2^n$ | $n^{\lg\lg n}$ | $e^n$ | $n^{1/\lg n}$ | $\lceil \lg n \rceil!$ |
| 3. | $2^{\sqrt{2\lg n}}$ | $(3/2)^n$ | $2$ | $\lg(n!)$ | $n$ | $\sqrt{\lg n}$ |
| 4. | $2^{(\lg n)^2}$ | $2^{2^n}$ | $n^2$ | $n \lg n$ | $(n+1)!$ | $4^{\lg n}$ |
| 5. | $\lg(\lg^* n)$ | $\lg^2 n$ | $(1 + \frac{1}{n})^n$ | $n^{\lg n}$ | $n!$ | $2^{(\lg n)/n}$ |
| 6. | $(\sqrt{2})^{\lg n}$ | $\lg^* n$ | $(n/\lg n)^2$ | $\sqrt{n})$ | $\lg^*(\lg n)$ | $1/n$ |

NOTE: to organize of this large list of functions, we ask that you first order each row. Then the rows are merged in pairs. Finally, perform a 3-way merge of the 3 lists. Show the intermediate lists of your computation (it allows us to visually verify your work). ♢

**Exercise 14.7:** Order the following functions:

$$n, \quad \lceil \lg n \rceil !, \quad \lceil \lg \lg n \rceil !, \quad n^{\lceil \lg \lg n \rceil !}, \quad 2^{\lg^* n}, \quad \lg^*(2^n), \quad \lg^*(\lg n), \quad \lg(\lg^* n).$$

♢

**Exercise 14.8:** (Purdom-Brown) Our summation rules already gives the $\Theta$-order of the summations below. This exercise is interested in sharper bounds:
(a) Show that $\sum_{i=1}^n i! = n![1 + \mathcal{O}(1/n)]$.
(b) $\sum_{i=1}^n 2^i \ln i = 2^{n+1}[\ln n - (1/n) + \mathcal{O}(n^{-2})]$. HINT: use $\ln i = \ln n - (i/n) + \mathcal{O}(i^2/n^2)$ for $i = 1, \ldots, n$. ♢

**Exercise 14.9:** (Knuth) What is the asymptotic behavior of $n^{1/n}$? of $n(n^{1/n} - 1)$?
HINT: take logs. Alternatively, expand $\prod_{i=1}^n e^{1/(in)}$. ♢

**Exercise 14.10:** Estimate the growth behavior of the solution to this recurrence: $T(n) = T(n/2)^2 + 1$. ♢

End Exercises

## §15. Summary of Chapter

This is a long chapter, so it is worthwhile giving a brief recap of the highlights.

1. Our goal is to solve recurrences for functions $T(n)$ that arise in analysis of algorithms. A key example is the Master Recurrence (112), $T(n) = aT(n/b) + d(n)$.
2. The two principles of our approach is to view all recurrences as real recurrences, and to solve them up to $\Theta$-order.
3. We understand "solving recurrences" to mean expressing the $\Theta$-order of $T(n)$ in terms of some familiar function $f(n)$, $T(n) = \Theta(f(n))$.
4. To do this, we need to find the $\Theta$-order of sums. For instance, the master recurrence reduces to a sum of the form $S_f(n) = \sum_{i \geqslant 0}^{\log_b n} d(b^i)$. Thus we must at least know how to convert such sums into familiar functions.
5. We introduces elementary and almost cookbook methods to solve such sums. The idea is to recognize sums as either polynomial- or exponential-types.
6. The rote method suffices for the Master Recurrence, but is helpless against the Multiterm Master Recurrence (123). So we introduce the method of Real Induction.

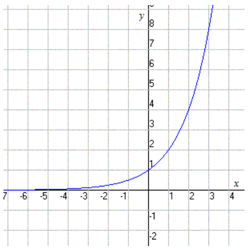
## §16 APPENDIX A: Exponential and Logarithm Functions

Next to the polynomials, the two most important functions in algorithmics are the **exponential function** and its inverse, the **logarithm function**. Many of our asymptotic results depend on their basic properties. For the student who wants to understand these properties, the following will guide them through some exercises. We define the **natural exponential function** to be

$$\exp(x) := \sum_{i=0}^{\infty} \frac{x^i}{i!}$$

for all real $x$. This definition is also good for complex $x$, but we do not need it. The **base of the natural logarithm** is defined to be the number

$$e := \exp(1) = \sum_{i=0}^{\infty} \frac{1}{i!} = 2.71828...$$



exponential function

The next Exercise derives some asymptotic properties of the exponential function.

**Exercise 16.1:** (a) $\exp(x)$ is continuous.
(b) $\frac{d\exp(x)}{dx} = \exp(x)$ and hence $\exp(x)$ has all derivatives.
(c) $\exp(x)$ is positive and strictly increasing.
(d) $\exp(x) \to 0$ as $x \to -\infty$, and $\exp(x) \to \infty$ as $x \to \infty$.
(e) $\exp(x+y) = \exp(x)\exp(y)$. ◇

We often need explicit bounds on exponential functions (not just its asymptotic behavior). Derive the following bounds:

**Exercise 16.2:**
(a) $\exp(x) \geqslant 1 + x$, with equality iff $x = 0$. Note that this holds for all $x$, even negative values; but this bound is trivial for $x \leqslant -1$.
(b) $\exp(x) > \frac{x^{n+1}}{(n+1)!}$ for $x > 0$. Hence $\exp(x)$ grow faster than any polynomial in $x$.
(c) For all real $n \geqslant 0$,

$$\left(1 + \frac{x}{n}\right)^n \leqslant e^x \leqslant \left(1 + \frac{x}{n}\right)^{n+(x/2)}.$$

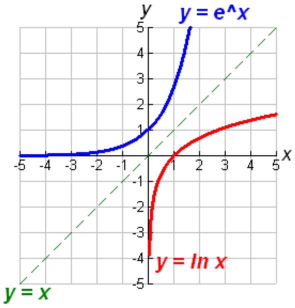
It follows that an alternative definition of $e^x$ is

$$e^x = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n.$$

(d) $\exp(x)\left(1 - \frac{x^2}{n}\right) \leqslant \left(1 + \frac{x}{n}\right)^n$ for all $x, n \in \mathbb{R}$, $n \geqslant 1$ and $|x| \leqslant n$. See [15]. In particular, $\exp(x) \leqslant 1/(1-x)$ for $|x| \leqslant 1$. ◇

The **natural logarithm** function $\ln(x)$ is the inverse of $\exp(x)$: $\ln(x)$ is defined[26] to be the real number $y$ such that $\exp(y) = x$. Since $\exp(y) > 0$, it follows that $\ln(x)$ is only defined for positive values of $x$. So $\ln(x)$ is a partial function which is defined iff $x > 0$.

**Exercise 16.3:** Show that
(a) $\frac{dy}{dx}\ln(x) = \frac{1}{x}$,
(b) $\ln(xy) = \ln(x) + \ln(y)$,
(c) $\ln(x)$ increases monotonically from $-\infty$ to $+\infty$ as $x$ increases from 0 to $+\infty$. ◇



graph of $\ln x$ (+ inverse $e^x$)

These two functions now allow us to define **general exponentiation** to any base $b > 1$: any real $\alpha$, we define

$$\exp_b(\alpha) := \exp(\alpha \ln(b)). \tag{147}$$

Usually, we write $\exp_b(\alpha)$ as $b^\alpha$. In particular, if $b = e$, this gives us the definition of $e^\alpha$, a familiar notation for $\exp(\alpha)$.

[26] This real value $y$ is called the principal value of the logarithm. That is because if we view $\exp(\cdot)$ as a complex function, then $\ln(x)$ is a multivalued function that takes all values of the form $y + 2n\pi$, $n \in \mathbb{Z}$.

We see from (147) that $b$ must be positive since $\ln(b)$ is otherwise undefined. Moreover, the case $b = 1$ is highly degenerate since $b^\alpha$ is identically equal to 1. It is easy to check that $(1/b)^\alpha = b^{-\alpha}$. Hence it is not necessary to consider exponentiation on bases $b$ that is less than 1: if $b < 1$, we can compute $b^\alpha$ by computing $B^{-\alpha}$ where $B = 1/b > 1$.

Once we have the definition of $exp_b(x) = b^x$, the **general logarithm** for any base $b \neq 1$ can be defined: $\log_b(x)$ is the inverse of the function $exp_b(x) = b^x$, i.e., $\log_b(x)$ is defined to be the $y$ such that $b^y = x$. Note that for $b > 1$, $\log_b(x)$ is well-defined for all $x > 0$.

*So $b^a$ and $\log_b a$ are derived from the "special cases" of $e^a$ and $\ln a$.*

**Why is the above definition of $b^x$ "correct"?** We have the straight forward relation that

$$2^{n+m} = 2^n \cdot 2^m$$

where $n, m$ are natural numbers. What if $n, m$ and 2 are replaced by real numbers $x, y$ and $b > 1$ (respectively)? Well, we would like identity

$$b^{x+y} = b^x \cdot b^y \tag{148}$$

to hold. The key observation is the fundamental identity

$$\sum_{i\geqslant 0} \frac{(x+y)^i}{i!} = \Big(\sum_{i\geqslant 0} \frac{x^i}{i!}\Big)\Big(\sum_{i\geqslant 0} \frac{y^i}{i!}\Big). \tag{149}$$

So, the formula (148) holds the special case $b = e$. Taking logs in (149), we get the fundamental relation

$$\ln X + \ln Y = \ln(XY)$$

where $X = e^x$ and $Y = e^y$. Similarly, we may verify that

$$b^{x+y} = b^x \cdot b^y.$$

and therefore $\log_b(XY) = \log_b X + \log_b Y$.

*Unless otherwise noted, the base $b$ of our general logarithm and exponentiation is assumed to satisfy $b > 1$.*

**Exercise 16.4:** We show some familiar properties: the base $b$ is omitted if it does not affect the stated property.

(a) The most basic properties are the following two:

$$\log(ab) = (\log a) + (\log b), \quad \log_b x = (\log_c x)/(\log_c b).$$

(b) $\log 1 = 0, \quad \log_b b = 1, \quad y = x^{\log_x y}, \quad \log(x^y) = y \log x.$

(c) $\log(1/x) = -\log x, \quad \log_b x = 1/(\log_x b), \quad a^{\log b} = b^{\log a}.$

(d) $\frac{d}{dx}(x^\alpha) = \alpha x^{\alpha-1}.$

(e) For $b > 1$, the function $\log_b(x)$ increases monotonically from $-\infty$ to $+\infty$ as $x$ increases from 0 to $\infty$. At the same time, for $0 < b < 1$, $\log_b(x)$ decreases monotonically from $+\infty$ to $-\infty$. ♢

**¶57. Varieties of logarithm and their notations.** When the actual value of the base $b$ of a logarithm is immaterial, we simple write 'log' without specifying the base. E.g., we write $\log(xy) = \log(x) + \log(y)$ without specifying the base. But it is important that the unspecified base is a fixed value $b > 1$. But there are three important bases: $b = e, b = 2, b = 10$, and we have a special notation for each: Clearly the natural logarithm $\ln x := \log_e x$ is the most important, all the other logarithms are defined in terms of it. But in computer science, we mainly use $\lg x := \log_2 x$. So

$\log x := \log_b x$

$\ln x := \log_e x$

$\lg x := \log_2 x$

$\lg x$ is often called the **Computer Science Logarithm**. Similarly, the **Engineers Logarithm** has 10 and is often denoted $\mathrm{Log} := \log_{10}$.

$\mathrm{Log}\, x := \log_{10} x$

We shall write $\log^{(k)} x$ for the $k$-fold application of the logarithm function to $x$. Thus $\log^{(2)} x = \log\log x$, and by definition, $\log^{(0)} x = x$. This is to be distinguished from "$\log^k n$" which equals $(\log n)^k$. On the black board, it is convenient to write $\ell\log n$ for $\log\log n$, and $\ell\ell\log n$ for $\log\log\log n$ (it does not pay to continue this process).

$\log^{(k)} x \neq \log^k x$

Finally, we have the **log-star function**. Starting from a value $x > 0$, we can keep taking logarithms until we get a value that is negative. If we can take logarithms at most $k$ times, then $\log^* x$ is defined to be $k$. By definition of log-star, if $k = \log^* x$ then $\log^{(k)} x \leqslant 0$ and $\log^{(k+1)} x$ is undefined. Notice that we have not specified the base of the logarithm. In most applications, the base of the log-star function is assumed to be 2. With this base, we see that $\log^*(x) = 0$ (resp., 1 and 2) iff $x \leqslant 0$ (resp., $0 < x \leqslant 1$ and $1 < x \leqslant 2$). So the range of log-star is[27] the set of natural numbers.

*$\log^* x$ is very, very slow growing*

There is another direction in the generalization of logarithms: we can define logarithms for negative numbers: using the rule $\ln(xy) = \ln x + \ln y$, it is sufficient to define $\ln(-1)$. The answer turns out to be $(2n+1)\pi\boldsymbol{i}$ where $\boldsymbol{i} = \sqrt{-1}$ and $n \in \mathbb{N}$. Two surprises here: we must go imaginary and give up single-valued functions. Once we go imaginary, we might as well allow arbitrary complex numbers $z = x + \boldsymbol{i}y$ as argument. Now $\ln : \mathbb{C} \to \mathbb{C}$. But $\ln 0$ remains undefined. This extension to complex numbers is not important for algorithmic analysis where we are interested in the growth rate of functions which reduces to the comparison of numbers. Unfortunately, complex numbers are not totally ordered like real numbers.

**¶58. Bounds on logarithms.** For approximations involving logarithms, it is useful to recall the Maclaurin series for logarithms:

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots = -\sum_{i=1}^{\infty} \frac{(-x)^i}{i} \tag{150}$$

valid for $-1 < x \leqslant 1$. Substituting $x$ by $-x$, we get

$$\ln(1-x) = -(x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots) = -\sum_{i=1}^{\infty} \frac{x^i}{i}, \qquad (-1 \leqslant x < 1).$$

E.g., this gives a lovely formula for $\ln 2$:

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = -\sum_{i=1}^{\infty} \frac{(-x)^i}{i}.$$

Furthermore, we obtain two very useful bounds:

$$x - \tfrac{x^2}{2} < \ln(1+x) < x \qquad (|x| < 1). \tag{151}$$

To see that $\ln(1+x) < x$ we must show that $R = \sum_{i=2}^{\infty} \frac{(-x)^i}{i} > 0$. This follows because if we pair up the terms in $R$ we obtain

$$R = \left(\tfrac{x^2}{2} - \tfrac{x^3}{3}\right) + \left(\tfrac{x^4}{4} - \tfrac{x^5}{5}\right) + \cdots,$$

which is clearly a sum of positive terms. A similar argument shows $\ln(1+x) > x - x^2/2$.

[27] We could have extended log-star to take all integer-values: $\log^*(x)$ is undefined for $x \leqslant 0$. For $0 < x < 1$, let $\log^*(x) := -k$ iff $k \geqslant 0$ is the number of times we must raise $x$ to the power of 2 until the result lies in the range $[1/2, 1)$.

> Have you ever wondered at the peculiar form (150)? Why give the series of $\ln(x)$ instead of $\ln(1+x)$? We must understand that the series expansion of a function $f(x)$ is typically the Taylor series expansion about the neighborhood of some chosen point $x = x_0$. The usual default is to choose $x_0 = 0$. The requested "series of $\ln(x)$" amounts to an expansion of $\ln(x)$ at $x = 0$. Unfortunately, $x = 0$ is a singularity for $\ln(x)$ function, and so no expansion is possible. The same phenomenon arises in the expansion of the square root function,
> $$\sqrt{1+x} = 1 + \tfrac{x}{2} - \tfrac{x^2}{8} + \tfrac{x^3}{16} - \cdots$$
> valid for $|x| < 1$ or $x = \pm 1$ (in general, $x$ is complex).

The formula (150) allows us to compute $\ln(y)$ for any $y \in (0, 2)$. How do we evaluate $\ln(y)$ for $y \geqslant 2$? Assume that we have good approximations to $\ln(2)$. Then we can write $y = 2^n(1+x)$ (i.e., $n$ is the number of times we must divide $y$ by 2 until its value is less than 2). Then we can evaluate $\ln(y)$ as $n\ln(2) + \ln(1+x)$. This procedure depends on having a good approximation to $\ln(2)$. Can we do this? One way is to use (150) with $x = 1$,

$$\ln 2 = \ln(1+1) = -\sum_{i\geqslant 1}(-1)^i/i = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots \tag{152}$$

(convergence of this series requires proof (150) assumes $|x| < 1$). But (152) converges very slowly, and the following alternative formula is much better:

$$\ln 2 = \sum_{k=1}^{\infty} \frac{1}{k2^k} \tag{153}$$

Using this rapidly converging series, we can quickly compute $\ln 2$ to any desired accuracy. To derive this series, note that $\frac{1}{1-x} = \sum_{i\geqslant 0} x^i$ and so $\int \frac{dx}{1-x} = \sum_{i\geqslant 0} x^{i+1}/(i+1) = \sum_{i\geqslant 1} x^i/i$. Putting $y = 1-x$, $\int \frac{dx}{1-x} = -\int \frac{dy}{y} = -\ln y = \ln(1/y)$. This shows

*Mother of Series again!*

$$\ln \frac{1}{1-x} = \sum_{i\geqslant 1} x^i/i,$$

and (153) is just the special case where $x = 1/2$.

Alternatively, to compute $\ln y$, we can write $y = n(1+x)$ where $n \in \mathbb{N}$ and write $\ln(y) = \ln(n) + \ln(1+x)$. E.g., we could pick $n$ to be $\lfloor y \rfloor$ or any integer larger than $y/2$. To evaluate $\ln(n)$ we use the fact $\ln(n) = H_n - \gamma - (2n)^{-1} - \mathcal{O}(n^{-2})$ (see §5). Of course, this method requires approximations Euler's constant $\gamma$ instead of $\ln 2$. Again, there are rapid approximations of $\gamma$.

Exercises

**Exercise 16.5:** (Open ended) Implement the above two suggested methods of computing $\ln x$ for arbitrary $x \in \mathbb{N}$. Evaluate their relative efficiency. ◇

End Exercises

# References

[1] J. L. Bentley, D. Haken, and J. B. Saxe. A general method for solving divide-and-conquer recurrences. *ACM SIGACT News*, 12(3):36–44, 1980.

[2] G. Dowek. Preliminary investigations on induction over real numbers, 2003. Manuscript, http://www.lix.polytechnique.fr/ dowek/publi.html.

[3] H. Edelsbrunner and E. Welzl. Halfplanar range search in linear space and $O(n^{0.695})$ query time. *Info. Processing Letters*, 23:289–293, 1986.

[4] M. H. Escardó and T. Streicher. Induction and recursion on the partial real line with applications to Real PCF. *Theoretical Computer Science*, 210(1):121–157, 1999.

[5] W. Feller. *An introduction to Probability Theory and its Applications.* Wiley, New York, 2nd edition edition, 1957. (Volumes 1 and 2).

[6] M. L. Fredman. *Growth Properties of a class of recursively defined functions.* PhD thesis, Stanford University, 1972. Technical Report No. STAN-CS-72-296. PhD Thesis.

[7] G. H. Gonnet. *Handbook of Algorithms and Data Structures.* International Computer Science Series. Addison-Wesley Publishing Company, London, 1984.

[8] D. H. Greene and D. E. Knuth. *Mathematics for the Analysis of Algorithms.* Birkhäuser, 2nd edition, 1982.

[9] G. H. Hardy. *Orders of Infinity.* Cambridge Tracts in Mathematics and Mathematical Physics, No. 12. Reprinted by Hafner Pub. Co., New York. Cambridge University Press, 1910.

[10] A. Karatsuba and Y. Ofman. Multiplication of many-digital numbers by automatic computers. *Doklady Akad. Nauk SSSR*, 145:293–294, 1962.

[11] D. E. Knuth. *The Art of Computer Programming: Fundamental Algorithms*, volume 1. Addison-Wesley, Boston, 2nd edition edition, 1975.

[12] G. S. Lueker. Some techniques for solving recurrences. *Computing Surveys*, 12(4):419–436, 1980.

[13] B. P. Mahony and I. J. Hayes. Using continuous real functions to model timed histories. In *Proc. 6th Australian Software Engineering Conf. (ASWEC91)*, pages 257–270. Australian Comp. Soc., 1991.

[14] B. Mishra and A. Siegel. (Class Lecture Notes) Analysis of Algorithms, January 28, 1991.

[15] D. S. Mitrinović. *Analytic Inequalities.* Springer-Verlag, New York, 1970.

[16] J. Paul Walton Purdom and C. A. Brown. *The Analysis of Algorithms.* Holt, Rinehart and Winston, New York, 1985.

[17] R. M. Verma. A general method and a master theorem for divide-and-conquer recurrences with applications. *J. Algorithms*, 16:67–79, 1994.

[18] X. Wang and Q. Fu. A frame for general divide-and-conquer recurrences. *Info. Processing Letters*, 59:45–51, 1996.

[19] C. K. Yap. A real elementary approach to the master recurrence and generalizations. In M. Ogihara and J. Tarui, editors, *8th Conf. on Theory and Applic. of Models of Computation (TAMC)*, pages 14–26. Springer, May 2011. LNCS No. 6648. May 23-25, Tokyo, Japan. The full paper may be obtained from the author's website.