

282 Math Preview

Chris Brown

June 29, 2011

Contents

1	Why This?	2
2	Logarithms	2
2.1	Basic Identities	2
2.2	Basic Consequences	3
3	Sums of Series	3
3.1	Arithmetic Series	3
3.2	Polynomial Series	4
3.3	Geometric Series	4
3.4	Harmonic Series	5
3.5	Arithmetic-Geometric Series	6
3.6	Using Integration	6
3.7	Who Knew?	8
4	Bounds	8
4.1	Big Bounds by Limits	8
4.2	Putting Limits to Work	8
4.3	Little Bounds by Limits	9
4.4	Properties	10
5	Simple Recurrence Techniques	10
5.1	Brute Force Substitution	11
5.2	Telescoping	11
5.3	The Master Method	13

1 Why This?

I've noticed that often the only math in an algorithms book is presented in one or two chapters and then is referred to for the next 1100 or so pages. So it's good to get this material nailed down early.

This is mostly reiteration of content from CLRS chapters 3 and 4, but I've used material from a couple of other Algorithms textbooks. This is meant to fill in some CLRS gaps, but I'm not sure the gaps are really there – maybe it's just been useful for me to pull this stuff together in one spot. If it helps anyone else, great, but it's certainly not meant to be any sort of self-contained tutorial.

Still, please report typos, bugs, thinkos, and unclarities, as well as suggestions for more content, to brown@cs.rochester.edu and maybe we can improve its chances of being useful.

2 Logarithms

2.1 Basic Identities

Definition: For $b > 1, x > 0$, $\log_b(x)$ is real L such that $b^L = x$.

Properties: (following easily from the definition)

- \log_b is a strictly increasing function: $x > y \Rightarrow \log_b(x) > \log_b(y)$
- \log_b is one-to-one: $\log_b(x) = \log_b(y) \Rightarrow x = y$.
- $\log_b(1) = 0$ for any b since $b^0 = 1$.
- $\log_b(b^a) = a$ (pretty much restates the definition).
- $\log_b(xy) = \log_b(x) + \log_b(y)$. If $b^L = x$ and $b^M = y$, then $b^L b^M = b^{L+M} = xy$ and by the previous result $\log_b(xy) = L + M = \log_b(x) + \log_b(y)$.
- $\log_b(x^a) = a \log_b(x)$. Use previous result a times with $y = x$.
- $x^{\log_b(y)} = y^{\log_b(x)}$. Use 2nd property above to justify taking logs on both sides: $\log_b(x) \log_b(y) = \log_b(y) \log_b(x)$.
- $\log_c(x) = \log_b(x) / \log_b(c)$. This important base-changing identity that establishes that logs to two different bases are related by a constant multiple. Let $L = \log_c(x)$, $M = \log_b(x)$, $N = \log_b(c)$. Then $c^L = x$, $b^M = x$, and $b^N = c$. But if $c = b^N$ then $x = c^L = (b^N)^L$, and also $x = b^M$, so $b^M = b^{NL} \Rightarrow M = NL \Rightarrow L = M/N$.

2.2 Basic Consequences

There are a number of little corollaries and techniques flowing from these identities that might cause some head-scratching when first seen. They are useful if you want to compare functions that are written using different bases. For instance, one can rewrite an exponentiation to a different base by first taking logs and then exponentiating to the base of the log, *comme ça*:

$$n^x = 2^{x \lg(n)},$$

or the more exciting

$$n^{\lg(n)} = 2^{\lg \lg(n)}.$$

or

$$n \lg(n) = 2^{\lg(n) + \lg \lg(n)} = 2^{1 + \frac{\lg \lg(n)}{\lg(n)}},$$

which if you're normal isn't obvious at first blush.

These examples do not address the complications introduced by arbitrary bases (*e.g.* e , 10, 3 ...) that would force the use of the base-changing identity, thus introducing a base-changing multiplier into the picture.

3 Sums of Series

3.1 Arithmetic Series

For instance

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

Why? Draw $n \times (n+1)$ rectangle and a jaggy diagonal and count the squares in one half. Or think of numbers 1 to n and pair them up first to last, 2nd to penultimate, etc. You get $n/2$ sums each equal to $n+1$. Questions: what if n is odd? Does this trick work for other summation limits?

These questions should lead you to think that maybe you can use the above counting tricks for linear variants on the sum. Indeed. Consider $2 + 5 + 8 + \dots + (3k-1)$. Rewritten as $3(1 + 2 + 3 + \dots + k) - (1 + 1 + 1 + \dots + 1)$ the sum is obviously $3(k(k+1)/2) - k$. OR you can again add the first and last, 2nd and next-last, etc. to get $k/2$ pairs of pair-sums each equal to $(3k+1)$, or $k(3k+1)/2$, which is the same answer, (*nicht?*).

3.2 Polynomial Series

The most familiar instance:

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

more generally,

$$\sum_{i=1}^n i^k \approx \frac{1}{k+1} n^{k+1}.$$

You might remember that the sum of the first n squares has a cubic term in the answer...turns out to be no accident. In general, summing over the exponentiation index:

$$\sum_{i=1}^n i^k \approx \frac{1}{k+1} n^{k+1}.$$

This is justified approximation to an integral (see later), but for any specific k (2 is a classroom favorite) the exact formula (involving a $k+1$ st degree polynomial) can be proved by induction.

3.3 Geometric Series

The old familiar

$$\sum_{i=0}^k 2^i = 2^{k+1} - 1$$

is of course visualized best by thinking of each term 2^i as a 1-bit in a binary number. Thus the sum = 1111...111 for $k+1$ bits, and if you add 1 you get 100000000 = 2^{k+1} .

Related is the special case

$$\sum_{i=0}^k \frac{1}{2^i} = 2 - \frac{1}{2^k}.$$

Bearing a close resemblance is the more general

$$\sum_{i=0}^k br^i = b \left(\frac{r^{k+1} - 1}{r - 1} \right),$$

with r often known as the *ratio*.

The last is a trivial generalization of

$$\sum_{i=0}^k r^i = \frac{r^{k+1} - 1}{r - 1},$$

which when $0 < r < 1$ leads to

$$\sum_{i=0}^k r^i \leq \frac{1}{1 - r}$$

as $k \rightarrow \infty$. In this tending-to-infinite case, the sum is quite easy to derive.

Let S be the infinite sum of powers. Then

$$S = 1 + r + r^2 + r^3 + r^4 + \dots$$

So

$$rS = r + r^2 + r^3 + r^4 + \dots,$$

and subtracting these two equations (only allowable if they are convergent), an infinite amount of right hand side vanishes to leave

$$S - rS = 1 \Rightarrow S = \frac{1}{1 - r}.$$

Cute, eh? This trick can generalize to more complex summand terms if one is careful (see the telescoping technique in Section 5.2).

3.4 Harmonic Series

We saw that the general case for polynomial series is

$$\sum_{i=1}^n i^k \approx \frac{1}{k+1} n^{k+1},$$

which does not work when $k = -1$. That case is the Harmonic series

$$H(n) = \sum_{i=1}^n \frac{1}{i} \approx \ln(n) + \gamma,$$

where $\gamma \approx 0.57721566\dots$ is *Euler's constant*.

3.5 Arithmetic-Geometric Series

In this example sum the index appears both as coefficient and exponent in the summand terms. Unfortunately our solution relies on the base being 2...

$$\sum_{i=1}^k i2^i = (k-1)2^{k+1} + 2.$$

The solution for this type of sum is an analog of integration by parts, which counts on producing sub-sums that cancel (except for first and last terms) and the sum of a leftover term that is easy to evaluate.

$$\sum_{i=1}^k i2^i = \sum_{i=1}^k i(2^{i+1} - 2^i),$$

(here is our reliance on base 2) – and now

$$= \sum_{i=1}^k i2^{i+1} - \sum_{i=0}^{k-1} (i+1)2^{i+1}.$$

We're aiming for a sum of terms of form $i2^{i+1}$ to emerge from the RHS, and sure enough...

$$= \sum_{i=1}^k i2^{i+1} - \sum_{i=0}^{k-1} i2^{i+1} - \sum_{i=0}^{k-1} 2^{i+1}$$

$$= k2^{k+1} - 0 - (2^{k+1} - 2) = (k-1)2^{k+1} + 2,$$

with a little care on the last sum.

3.6 Using Integration

With some simple ideas rendered mathematically precise we can translate results from continuous mathematics into results for the discrete sums we use.

One simple idea is *monotonic* (nondecreasing) (e.g. $\log(x)$, x^2 , 2^x if $x > 0$; or the discontinuous $\lfloor x \rfloor$), and *antimonotonic* (non-increasing) functions, (e.g. $1/x$).

Convex functions are those that “never curve downward” (like x , $1/x$, e^x , x^4). These puppies have always looked concave to me; I have to remember “convex down”. A function can be convex but not monotonic, or could be monotonic but not convex. Clearly \sqrt{x} is not convex, nor is $\log(x)$. A discontinuous function cannot be convex. Convexity can be proved by showing, for a real function, that the average value of the function at two points is above the function of their average. Convexity for a function on the integers can be proved by showing the same thing for all adjacent sets of three integers: $f(n+1)$ is at most the average of $f(n)$ and $f(n+2)$.

An integer function $f(n)$ can be extended to a real function $f^*(x)$ simply by linear interpolation. Then we have some useful properties:

1. $f(n)$ is monotonic (convex) $\Leftrightarrow f^*(x)$ is monotonic (convex).
2. $f'(x)$ (1st derivative) exists and is nonnegative $\Rightarrow f(x)$ monotonic.
3. $f'(x)$ exists and is monotonic $\Rightarrow f(x)$ convex.
4. Thus $f''(x)$ exists and is nonnegative $\Rightarrow f(x)$ convex.

See CLRS Appendix 2 on integral bounds and nice pictures explaining why they work... the idea is to put upper and lower bounds on discrete sums by pairs of definite integrals on the corresponding functions.

Some useful integration formulae (the *Theory Math Cheat Sheet* linked off the main course page has many more):

$$\int_0^n x^k dx = \frac{1}{k+1} n^{k+1}$$

$$\int_0^n e^{ax} dx = \frac{1}{a}(e^{an} - 1)$$

$$\int_1^n x^k \ln(x) dx = \frac{1}{k+1} n^{k+1} \ln(n) - \frac{1}{(k+1)^2} n^{k+1}.$$

CLRS on p. 1067 show a bound for the harmonic series (Section 3.4). Another example is to use a simple case of the last formula to get a lower bound for the sum of logarithms — which is what? The logarithm of a factorial!

$$\lg(n!) = \sum_{i=1}^n \lg i = 0 + \sum_{i=2}^n \lg i \geq \int_1^n \lg x dx$$

by CRLS formula (A.11).

Changing bases,

$$\begin{aligned} \int_1^n \lg x dx &= \int_1^n \lg(e) \ln x dx = \lg(e) \int_1^n \ln x dx \\ &= (\lg e)(x \ln x - x) \Big|_1^n = (\lg e)(n \ln n - n + 1) \\ &= n \lg n - n \lg e + \lg e \geq n \lg n - n \lg e. \end{aligned}$$

$\lg e \leq 1.443$, so

$$\sum_{i=1}^n \lg i \geq n \lg n - 1.443n$$

3.7 Who Knew?

It seems that the last example is related to the derivation of the very useful *Stirling's Formula*, which bounds $n!$.

$$\left(\frac{n}{e}\right)^n \sqrt{2\pi n} \leq n! \leq \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \left(1 + \frac{1}{11n}\right),$$

(for $n \geq 1$).

4 Bounds

4.1 Big Bounds by Limits

Not much to add on bounds. Besides the traditional “there’s a c and n_0 such that for $n > n_0$ *etc. etc.*,” one can use infinite limits to determine the order of functions, and doing so is often easy because of L’Hôpital’s rule, so here’s the idea.

A function $f \in O(g)$ if:

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = c < \infty,$$

for nonnegative c (including the case $c = 0$.) So if the limit exists and is not infinite, f grows no faster than g and so $f \in O(g)$.

If the limit is ∞ then f grows faster than g (see below).

A function $f \in \Omega(g)$ if:

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} > 0,$$

including the case that the limit is ∞ .

This leaves:

A function $f \in \Theta(g)$ if:

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = c$$

for some constant $0 < c < \infty$.

4.2 Putting Limits to Work

L’Hôpital’s Rule: If f and g are differentiable functions with derivatives f' and g' , and

$$\lim_{n \rightarrow \infty} f(n) = \lim_{n \rightarrow \infty} g(n) = \infty,$$

then

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

So let's check what happens with some favorite functions like $f(n) = n^2$ and $g(n) = n \lg n$. We expect that $f \notin O(g)$ but $g \in O(f)$. In the ratio of functions a factor of n drops out so we're interested in

$$\lim_{n \rightarrow \infty} \frac{n}{\lg(n)}.$$

But we don't like differentiating $\lg(n)$, so we change base by the last identity in Section 2 to get the more friendly \ln function: $\lg(n) = \ln(n)/\ln(2)$. So by L'Hôpital's Rule,

$$\lim_{n \rightarrow \infty} \frac{n \ln(2)}{\ln(n)} = \lim_{n \rightarrow \infty} \frac{\ln(2)}{(1/n)} = \lim_{n \rightarrow \infty} n \ln(2) = \infty.$$

Since the ∞ case isn't allowed in the $O(n)$ test, f is not $O(g)$, but since the limit of the ratio $g(n)/f(n)$ must go to 0, g is $O(f)$.

4.3 Little Bounds by Limits

For strictly smaller and greater growth rates we have the “little oh” and “little omega” concepts. You're probably way ahead of me, since the limit forms of these definitions are obvious:

With f and g functions from non-negative integers into the positive reals,

$f \in o(g)$ is the set of functions f such that

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0,$$

and

$f \in \omega(g)$ is the set of functions f such that

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

It's easy to remember that “little oh” functions are the smaller functions in “Big Oh”. Less intuitively, the “little omega” functions are the LARGER ones in “Big Omega”.

4.4 Properties

CLRS have lots of fun exercises so you can prove properties of these bounding functions.

For instance:

1. Membership in $o, \omega, O, \Omega, \Theta$ are (each) transitive.
2. $f \in O(g) \Leftrightarrow g \in \Omega(f)$.
3. $f \in \Theta(g) \Rightarrow g \in \Theta(f)$.
4. Θ defines an equivalence relation whose classes are called complexity classes.
5. $O(f + g) = O(\max(f, g))$, with similar equations for Θ, Ω .
6. Our buddy L'Hôpital lets us prove: $\lg(n) \in o(n^a), a > 0$. So the log grows more slowly than any positive power of n , like the .000001 power.
7. Similarly, $n^k \in o(2^n), k > 0$. So powers of n grow more slowly than 2^n (in fact more slowly than any b^n where $b > 1$.)

The asymptotic order of some common summations is easy to determine. CLRS has Appendix A.2 entirely devoted to this topic. I don't know how the following fits into what it says, but does have some good examples of calculating bounds and some important caveats, including a neat "proof" that the sum of the first n integers is $O(n)$.

So some summation factoids follow:

If d is a nonnegative constant and $r \neq 1$ be a positive constant,

The sum of a polynomial series $\sum_{i=1}^n i^d \in \Theta(n^{d+1})$.

The sum of a geometric series $\sum_{i=a}^b r^i \in \Theta$ of its largest term. Remember we disallow $r = 1$, and note that usually b is some function of the problem size n .

The sum of a logarithmic series $\sum_{i=1}^n \log(i) \in \Theta(n \log(n))$.

The sum of a polynomial-logarithmic series $\sum_{i=1}^n i^d \log(i) \in \Theta(n^{d+1} \log(n))$.

The proof for the geometric series follows from the formula for its sum (Section 3.3) and the proofs for the other cases rest on the fact that the functions are monotonic and can be bounded by "upper" and "lower" rectangles. *E.g.* for n^d use the upper rectangle $[(0, 0), (n, 0), (n, n^d), (0, n^d)]$ and the lower rectangle $[(n/2, 0), (n, 0), (n, (n^d)/(2^d)), (n/2, (n^d)/(2^d))]$. It turns out these upper-bound rectangles and lower-bound rectangles both grow at the same rate, hence the function has to grow at that rate too. A picture would be good but I'm too lazy.

5 Simple Recurrence Techniques

CLRS presents some sophisticated techniques for solving recurrences but deemphasizes a couple of simple ones.

5.1 Brute Force Substitution

Probably the most obvious thing to do with a recurrence is just to use it to rewrite itself.

In the canonical mergesort example, start with the original

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

So what is $2T(n/2)$? We simply substitute $n/2$ in the equation and get

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

so

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

Plodding onwards,

$$4T(n/4) = 4(2T(n/8)) + n/4 = 8T(n/8) + n,$$

so

$$T(n) = 8T(n/8) + 3n.$$

So we see (and can prove by induction if we wish)

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

We know there will only be $k = \lg n$ terms, so

$$T(n) = nT(1) + n \lg n = n \lg n + n = O(n \log n).$$

The ugliness of those right hand side partial results is the target of the next technique, which with a little foresight aims to make the rewriting trivial.

5.2 Telescoping

The *telescoping technique* is mentioned in CLRS Appendix A1. The crucial observation is that for any sequence $a_0, a_1, a_2, \dots, a_n$,

$$\sum_{i=1}^n (a_i - a_{i-1}) = a_n - a_0.$$

We used this idea for our infinite-series-solving trick in Section 3.3.

Continuing with the mergesort example:

$$T(1) = 1$$

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

but it is not in the correct form to telescope. We want the right hand use of $T(i)$ to have exactly the same form as the left hand side $T(i - 1)$. We notice the pesky 2 and also the pesky n messing up the right hand side.... We want the RHS to be a function of $n/2$. Hmmmm. Inspiration strikes and we divide both sides by n to obtain

$$\frac{T(n)}{n} = \frac{T(n/2)}{n/2} + 1.$$

Fair warning.... this sort of pre-adjustment is an aspect of the telescoping method that is a bit creative and varies from recurrence to recurrence.

Anyway now we're set. Our equation now works for any n that is a power of two (removing this assumption is possible and gives almost the same answer) and we can keep rewriting thus:

$$\frac{T(n/2)}{n/2} = \frac{T(n/4)}{n/4} + 1,$$

$$\frac{T(n/4)}{n/4} = \frac{T(n/8)}{n/8} + 1,$$

down to

$$\frac{T(2)}{2} = \frac{T(1)}{1} + 1.$$

Finally, we can add up all these equations. The left hand side is our original sum, and there are canceling pairs of terms for all but the first on the LHS and the last on the RHS. There are $\lg(n)$ rewrites since we divide n by 2 every time, so

$$\frac{T(n)}{n} = \frac{T(1)}{1} + \lg n,$$

and multiplying by n gives us

$$T(n) = n \lg n + n = O(n \log n).$$

Now the idea of adding up all the equations is nice conceptually, but in practice we usually see telescoping presented and used just as re-writing a RHS term: above, just rewriting $\frac{T(n/2)}{n/2}$ as $\frac{T(n/4)}{n/4} + 1$, say.

5.3 The Master Method

The master method (MM) (a.k.a. the Master Theorem) is for solving recurrences of the form

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

Which is best visualized by a recursion tree that splits up the problem into a smaller versions (smaller by the factor b) and puts together subproblems at the current level at a cost of $f(n)$. The sum of non-recursive costs (the $f(n)$ s) at every level is the *row-sum*.

CLRS seems to rely on their MM proof for intuition as to why it works. They spend a fair amount of time and space on it (their Section 4.4, with some helpful graphics), and some books have even hairier proofs. Here I just want to make sure I'm understanding what's going on. The two big questions seem to be: In CLRS Theorem 4.1 (their Section 4.3):

1. What do those three cases mean?
2. Where did that $\log_b a$ come from?

OK. The problem size drops by factor of b for every unit depth increase. Thus we hit the leaves of the tree at depth D such that $(n/b^D) = 1$. Rearranging and taking logs yields $D = \lg(n)/\lg(b)$, which is in $\Theta(\log n)$.

What the MM is telling us, in fact the answer to 1), is that the row-sums are not the same at all depths. Hold that thought.

Let's figure out how many leaves the tree has. With branching factor a , the number of nodes at depth D is $L = a^D$. Taking logs gives $\lg(L) = D \lg a = (\lg(a)/\lg(b)) \lg(n)$, (using our expression for D from derived just above). The ratio $E = (\lg(a)/\lg(b))$ relating these tree properties is important. In fact it is the answer to question 2), since by the last property of logs in Section 2, (the base-changing formula), $E = \log_b(a)$. Call E the "critical exponent".

Thus the number of leaves in the recursion tree is about $L = n^E$. Also we know the depth of the tree (the number of row-sums) is about $D = \lg(n)/\lg(b)$. The row-sum for the row of leaves at depth D is $\Theta(n^E)$, or simply n^E if the base case cost is 1.

Last, the value $T(n)$ we are looking for is the sum of the non-recursive costs of all the nodes in the tree, which is the sum of the row-sums.

Often the row-sums form a geometric series, or can be approximated (closely bounded) by geometric series (here, a series of the form $\sum_{d=0}^D ar^d$ — see Section 3.3.) The ratio r is important in determining how the series acts. In particular, we know (Section 4) that if $r \neq 1$, the sum is in Θ of its largest term. From what we know and that fact we can deduce the following *Little Master Theorem*:

1. Roughly the $r > 1$ case. If the row-sums form an increasing geometric series (with row 0 at the root of the tree), then $T(n) \in \Theta(n^E)$, with E the critical exponent (a.k.a. $\log_b(a)$). The cost is proportional to the number of *leaves* in the recursion tree (since the sum is dominated by the largest member of the series, the last term).

2. Roughly the $r = 1$ case. If the row-sums are about constant, $T(n) \in \Theta(f(n) \log(n))$ (since there are $\Theta(\log(n))$ equal terms.)
3. Roughly the $r < 1$ case. If the row-sums are a decreasing geometric series, then $T(n) \in \Theta(f(n))$; that is, all the real work is done at the root in putting together the subproblems — the sum is dominated by the first term.

That's the intuition behind the MM in terms of geometric series, which are simple and easy to believe. The MM is a more general result with the same intuitions. The graphics in CLRS are quite helpful I think. If you don't like CLRS's proof, there's a nice long one in Johnsonbaugh and Schaefer.