Toolkit
- ~odonnell/toolkit13/
- http://piazza.com/cmu/fall2013/15859t
- signup sheet
- scribe → [[name a scribe]]

books: · De Bruijn
· Graham-Knuth-Patashnik

Lecture 1 – Asymptotics

[[We all know...]] $\sum_{i=1}^{n} i = \frac{n(n+1)}{2} = \frac{1}{2}n^2 + \frac{1}{2}n = O(n^2)$

rec: " $f(x) = O(g(x)) \quad (x \to \infty)$ " ← [[often omitted, esp. if "n"]]

$\equiv \exists C, x_0$ s.t. $|f(x)| \le C \cdot g(x) \quad \forall x \ge x_0$

THIS CLASS: $0 \le f(x)$ [[I like it to be only used for eventually nonneg things.]]

$(x \to 0^+)$ ← [[often omitted if called "ε"]]

(— " —) $\forall 0 \le x \le x_0$

ext: [[not totally std., but also in this class:]]

$O(g(x)) \equiv$ an anon. fcn $f(x)$ s.t. $0 \le f(x) \le Cg(x) \quad \forall x \ge x_0$.

e.g. $\sum_{i=1}^{n} i = \frac{1}{2}n^2 + O(n)$. [[This is better/preferred]]

$= \frac{1}{2}n^2\left(1 + O\left(\frac{1}{n}\right)\right)$ [[shows "multiplic. error"]]

[[analogy:]] O $\le$, Ω $\ge$, Θ $=$

$f(x) = \Omega(g(x))$: $\exists c > 0, x_0$ s.t. $f(x) \ge c\,g(x) \quad \forall x \ge x_0$.

o $<$, ω $>$: $f(x) = o(g(x))$: $\frac{f(x)}{g(x)} \to 0$ as $x \to \infty$

[[one more]]: $f(x) \sim g(x)$ [[\sim]] : $\frac{f(x)}{g(x)} \to 1$.
[[stronger than Θ]]

e.g., $\sum_{i=1}^{n} i = \Theta(n^2), \sim \frac{1}{2}n^2$.

(2)

- "$f(x) \le \mathrm{poly}(g(x))$" $\equiv$ $f(x) = g(x)^{O(1)}$.
- $f(x) = \tilde{O}(g(x))$, $x \to \infty$: $f(x) = g(x) \cdot \mathrm{poly}(\log g(x))$
  ["lazy" version] Also applies to Ω, Θ]]

  e.g., $n^2 \log^3 n = \tilde{O}(n^2)$. $n^5 3^n = \tilde{O}(3^n)$. $\stackrel{?}{=} \tilde{O}(2^n)$ [no!]]

  if $x \to 0$: $f(x) = g(x)\, \mathrm{polylog}(\frac{1}{x})$.

  e.g. $\varepsilon \log^2(\frac{1}{\varepsilon}) = \tilde{O}(\varepsilon)$.

[[Well-known example, from Coupon Collector... Harmonic #]]

$$H_n := \sum_{i=1}^{n} \frac{1}{i} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n}.$$

[[Probably know it's $\sim \ln n$...]]

[[start elementary]]

$$\le \underbrace{1}_{1} + \frac{1}{2} + \frac{1}{2} + \underbrace{\frac{1}{4} + \cdots \frac{1}{4}}_{1} + \frac{1}{8} \text{ --- } + \quad \le \lfloor \log_2 n \rfloor + 1 \quad \sim \log_2 n$$

[[little thought re edge case]]

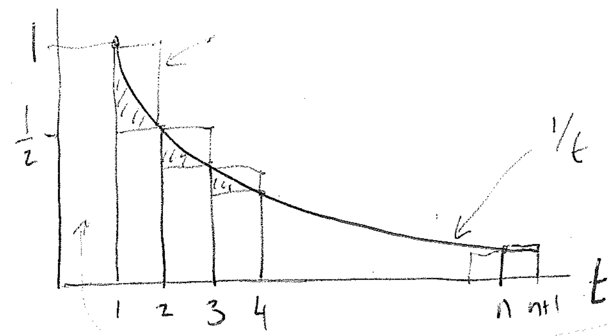
$$\ge 1 + \underbrace{\frac{1}{2}}_{1/2} + \underbrace{\frac{1}{4} + \frac{1}{4}}_{1/2} + \underbrace{\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}}_{1/2} + \frac{1}{10} + \cdots$$

$$\ge \frac{1}{2} \lceil \log_2 n \rceil + 1 \quad \sim \frac{1}{2} \log_2 n$$

$\therefore H_n = \Theta(\log n)$

[[often all you need]]

[[Rarely used (IMO) integral trick lets you improve this]]



[upper rects]

$$\therefore\ H_n \geq \int_1^{n+1} \frac{dt}{t} = \ln(n+1) \geq \ln n.$$

[lower]

$$H_n \leq 1 + \int_1^n \frac{dt}{t} = \ln n + 1$$

[[so we have it to w/i ± 1]]

[[~~mistake~~ error region is $\approx \frac{1}{2}$, must cvg. to some #:]]

[sidebar] $\ln(n+1)$?
["extract main term" method]

$$= \ln\left[n\left(1+\tfrac{1}{n}\right)\right]$$
$$= \ln n + \ln\left(1+\tfrac{1}{n}\right) = \ln n + \Theta\left(\tfrac{1}{n}\right)$$

$$H_n = \ln n + \gamma \underline{\pm\ O\left(\tfrac{1}{n}\right)}$$

[Euler–Mascheroni] $\approx .577$]

[why? :)]

[asympt(sum(1/i, i=1..n), n) in Maple :)]

[Taylor]

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots \qquad (-1 < x \leq 0)$$

$$\sim x \qquad (-O(x^2)) \qquad (x \to 0)$$

[inverse of "perhaps THE most usefull approx. ever"]

$$e^x = 1 + x \underbrace{+\ O(x^2)}_{\text{in } [0,\ x^2]\ \forall -1 \leq x \leq 1,\ \text{say.}} , \quad x \to 0$$

$$\left(\because\ e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \dots\right)$$

4

"Birthday paradox" [[abt. 23 ppl in room before $\approx 1/2$ chance of duplicate bday. Comes up all the time in TCS, goes under the B'day name, sillily enough]]

[[general ver:]] m bins, n balls $\rightarrow$ indep. thrown in rand bin.
[[365]]

$$p_{n,m} := \Pr[\text{no "collision"}] = 1 \cdot \left(1-\frac{1}{m}\right)\left(1-\frac{2}{m}\right)\cdots\left(1-\frac{n-1}{m}\right).$$

[[all balls in diff bins]] $(n \le m)$

[[When you see such a product, you should usually take its ln. But to illustrate $e^x \approx 1+x$ we'll take dir approach:]]

$\boxed{\exp(t) \equiv e^t}$

$$p_{n,m} \le \exp(0)\exp\left(-\frac{1}{m}\right)\exp\left(-\frac{2}{m}\right)\cdots\exp\left(-\frac{n-1}{m}\right)$$

$$= \underline{\exp\left(-\frac{n(n-1)}{2m}\right)}$$

[[This is highly accurate]]

$p_{n,m} \geq$?

$$\left(1-\frac{i}{m}\right) \overset{ex}{\geq} \exp\left(1-\frac{i}{m} - O\left(\frac{i^2}{m^2}\right)\right)$$

$$\geq \exp\left(-\frac{n(n-1)}{2m}\right) \underbrace{\exp\left(-\left(O\left(\frac{1}{m^2}\right)+\cdots+O\left(\frac{n^2}{m^2}\right)\right)\right)}$$

$$\exp\left(-O(n^3)/m^2\right)$$

[[why? $1+2^2+\cdots+n^2 = O(n^3)$]]

$$\geq 1 - \frac{O(n^3)}{m^2}$$

$$\therefore \quad p_{n,m} = \exp\left(-\frac{n(n-1)}{2m}\right)\cdot\left(1 - O(n^3/m^2)\right)$$

[[n-1 annoying]]

$$= \exp\left(-\frac{n^2}{2m}\right) \underbrace{\exp\left(\frac{n}{2m}\right)}_{\le 1 + O(n/m)} \left(1 - O(n^3/m^2)\right).$$

[[Which error bigger? $\frac{n}{m} = \frac{n^3}{m^2}$ $\Leftrightarrow$ $n = \sqrt{m}$. Former if $n \le \sqrt{m}$]]

$= \frac{1}{2}$ if $\frac{n^2}{2m} = \ln 2$,

$\Leftarrow n = \sqrt{2 \ln 2 \, m}$ [[not actually an int]]

$= \boxed{\Theta(\sqrt{m})}$ [[remember]].

[[$n = \sqrt{2\ln 2}\sqrt{m} \pm O(1)$

$\Rightarrow p_{n,m} = \frac{1}{2} \pm O(\frac{1}{\sqrt{m}})$]].

[[$n!$ & Stirling]]

$n! = n(n-1)(n-2) \cdots 2 \cdot 1$ [[? Start simple!]]

$\le n \; n \; n \; \cdots \le n^n$ [[n^{n-1}]]

$\ge \underbrace{(n/2)(n/2)\cdots(n/2)}_{n/2 \text{ times}}$ [[too lazy to look at edge case]]

$\ge \left(\frac{n}{2}\right)^{n/2}$

[[How do these compare? If things look rough, & you have expons, take logs!]]

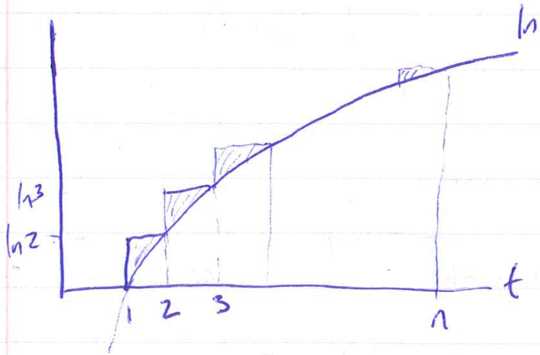
$\therefore \; \frac{n}{2}\ln\left(\frac{n}{2}\right) \le \ln(n!) \le n \ln n$

$= \frac{1}{2} n \ln n - \Theta(n)$ [[so we got it up to factor $\frac{1}{2}$ — in exponent]].

$\therefore \; n! = 2^{\Theta(n \log n)}$. [[let's do better]].

(6)

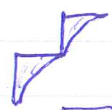
$\ln(n!) = \ln 1 + \ln 2 + \ln 3 + \cdots + \ln n.$ [[Time for integral trick]]



$\therefore \ln(n!) \geq \int_1^n \ln t \, dt$

$\overset{(ex)}{=} t\ln t - t \Big]_1^n$

$= n\ln n - n + 1.$

Upper bound: add in 

straighten:  $\Big\}\ln n = \frac{1}{2}\ln n.$

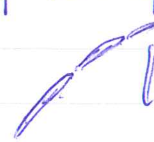
$$\therefore \quad n\ln n - n + 1 \leq \overset{\ln(n!)}{} \leq n\ln n - n + \tfrac{1}{2}\ln n + 1$$

$$\Rightarrow \quad n^n \cdot e^{-n} \cdot e \leq n! \leq n^n e^{-n} \sqrt{n} \cdot e$$ [[great! 2nd order decrease]]

$$\overset{\|}{\tilde{\Theta}}\left(\left(\tfrac{n}{e}\right)^n\right).$$

[[Which is it? $\cdot\sqrt{n}$? $\cdot\,\Theta(1)$? (Clearly $\sqrt{n}$ est. is better.]]

Gap from upper bd?

 [slivers]. Ex: area of slivers is $O(1)$. $(\leq \frac{1}{8})$

$\therefore \quad n! \geq n\ln n - n + \frac{1}{2}\ln n - O(1)$

$\therefore \quad n! = \Theta\left(\left(\frac{n}{e}\right)^n \cdot \sqrt{n}\right).$ [[Rarely need more]]

[[getting const. is like Euler–Maschroni]]

Stirling's Formula: $n! \sim \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n.$ $\left(= \cdots \cdot \left(1 \pm O\left(\frac{1}{n}\right)\right)\right).$

[[We'll understand why the $\sqrt{2\pi}$ when we study Gaussians.]]

[[Binomial coeffs]]: $\binom{n}{k}$. $\underline{k \text{ small}}$ e.g. $\binom{n}{3} = \frac{n(n-1)(n-2)}{3!} \sim \frac{n^3}{6}$.

[[take out main terms]]

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n^k}{k!} \cdot \underbrace{1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)}$$

[[like in birthday paradox!]]

$$\approx 1 - \frac{k^2}{2n} \text{ if } k \text{ small}$$

$$= (1-o_n(1)) \text{ if } k = o(\sqrt{n}).$$

$$\therefore \sim \frac{n^k}{k!} \quad \text{for } k = o(\sqrt{n}).$$

In general: $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n!}{(pn)!(qn)!}$ for $p = \frac{k}{n}$, $q = 1-p$ [[prob. notation]]

$$\underset{\text{stirling}}{\sim} \frac{\sqrt{2\pi n}\,(n/e)^n}{\sqrt{2\pi pn}\,(pn/e)^{pn}\sqrt{2\pi qn}\,\left(\frac{qn}{e}\right)^{qn}} \frac{\left(1 \pm O\left(\frac{1}{n}\right)\right)}{\left(1 \pm O\left(\frac{1}{pn}\right)\right)\left(1 \pm O\left(\frac{1}{qn}\right)\right)}$$

[[ignore for a sec.]]

$$= \frac{1}{\sqrt{2\pi n}} \frac{1}{\sqrt{pq}} \left(\frac{1}{p^{pn} q^{qn}}\right)^n$$

$$= \frac{1}{\sqrt{2\pi pq}} \frac{1}{\sqrt{n}} \cdot 2^{H(p)n}, \quad \leftarrow = 2^{\Theta(n)} \text{ for "const." } p$$

$$H(p) = \text{"binary entropy"} = p\log_2\left(\frac{1}{p}\right) + q\log_2\left(\frac{1}{q}\right)$$

error factor: $1 \pm O\left(\frac{1}{\sqrt{n}}\right)$ if $k, n-k \geq \sqrt{n}$.

$1 \pm O\left(\frac{1}{n}\right)$ if $.01 \leq p, q \leq .99$ [[quite important]]

e.g. $\underset{(p=q=1/2)}{\binom{n}{n/2}} = \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sqrt{n}} \cdot 2^n$ $\therefore$ $\Pr[\text{exactly } \frac{n}{2} \text{ H when flipping } n \text{ coins}] \sim \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sqrt{n}} = \Theta\left(\frac{1}{\sqrt{n}}\right)$