T1 – good average, tests in the mail... soon!
A1 – also good, also soon
- too many tests? François talking to A+S
- annotations - you touched my heart!

CSC236 fall 2016

more complexity: mergesort

Danny Heap

heap@cs.toronto.edu / BA4270 (behind elevators)

http://www.cdf.toronto.edu/~heap/236/F16/

416-978-5899

Using Introduction to the Theory of Computation, Chapter 3

Outline

vexing complexity

mergesort

Divide-and-conquer

Notes

Upper bound on $T(n)$ - binary search

trouble!

$$T(n) = \begin{cases} 1 & n = 1 \\ 1 + \max\left(T(\lceil n/2 \rceil), T(\lfloor n/2 \rfloor)\right) \end{cases}$$

- assume (not hard to prove) T is monotonic
- worst case always uses $T(\lceil \frac{n}{2} \rceil)$

WTP $T(n) \leq c \lg n$

problem

$T(n) \leq c \lg(n-1) \leq c \lg n$

Pick $c = \underline{3}$, then $c \in \mathbb{R}^+$

Pick $B = \underline{3}$, then $B \in \mathbb{N}$

Let $n \in \mathbb{N}$, $n \geq B$. assume $T(i) \leq c \lg(i-1)\ \forall\ B \leq i < n$.

~~Case $n \geq 5$~~

$$\begin{aligned} T(n) &= 1 + T(\lceil \tfrac{n}{2} \rceil) && \text{\# by def + } T \text{ is monotonic} \\ &\leq 1 + c \lg(\lceil \tfrac{n}{2} \rceil - 1) && \text{\# by IH, } B \leq \lceil \tfrac{n}{2} \rceil < n \\ &\leq 1 + c \lg(\tfrac{n+1}{2} - 1) && \text{\# } \tfrac{n+1}{2} \geq \lceil n/2 \rceil \\ &= 1 + c \lg(\tfrac{n-1}{2}) \\ &= 1 + c(\lg(n-1) - \lg 2) \\ &= 1 + c \lg(n-1) - c \\ &= 1 - c + c \lg(n-1) && \text{\# } c \geq 1 \\ &\leq c \lg(n-1) \end{aligned}$$

base cases 3 + 4

$n = 0$ work?

$n = 1$ work?

- not defined

$n = 2$

$\cancel{T(2) = 2 \leq c \lg(2-1)}$

$T(3) = 3 \leq c \lg(3-1)$ ✓

Base cases: $n = 3, 4$

recurrence for MergeSort

```
MergeSort(A,b,e):
    if b == e:  return
    m = (b + e) / 2
    MergeSort(A,b,m)
    MergeSort(A,m+1,e)
    # merge sorted A[b..m] and A[m+1..e] back into A[b..e]
    for i in [b,...,e]:  B[i] = A[i]
    c = b
    d = m+1
    for in [b,...,e]:
        if d > e or (c <= m and B[c] < B[d]):
            A[i] = B[c]
            c = c + 1
        else:  # d <= e and (c > m or B[c] >= B[d])
            A[i] = B[d]
            d = d + 1
```

c_1

c_2

$T(\lceil \frac{n}{2} \rceil)$

$T(\lfloor \frac{n}{2} \rfloor)$

$c_3 \cdot n$

constant

$c_4 \cdot n$

constant.

save arith by neglect,

$$T(n) = \begin{cases} c & n = 1 \\ T(\lceil \frac{n}{2} \rceil) + T(\lfloor \frac{n}{2} \rfloor) + n & n > 1 \end{cases}$$

$c'n$

choose my time units so that $c' = 1$

$\boxed{\text{Unwind}}$ (repeated substitution)

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

What happens when $n = 2^k$, some $k \in \mathbb{N}$?

$$T(n) = T(2^k) = 2^1 T(2^{k-1}) + 2^k$$

$$= 2^1(2^1 T(2^{k-2}) + 2^{k-1}) + 2^k$$

$$= 2^2 T(2^{k-2}) + 2 \cdot 2^k$$

$$= 2^2(2^1 T(2^{k-3}) + 2^{k-2}) + 2 \cdot 2^k.$$

$$= 2^3 T(2^{k-3}) + 3 \cdot 2^k$$

$\vdots$ intuition happens!!

prove by induction on k

$n = 2^k$

so $\lfloor n/2 \rfloor = \lceil n/2 \rceil$

$$= 2^k T(2^{k-k}) + k \cdot 2^k$$

$$= nc + n \cdot \lg n = n \lg n + c \cdot n$$

Prove that T is non-decreasing

$$T = \begin{cases} c & n = 1 \\ T(\lceil \frac{n}{2} \rceil) + T(\lfloor n/2 \rfloor) + n \end{cases}$$

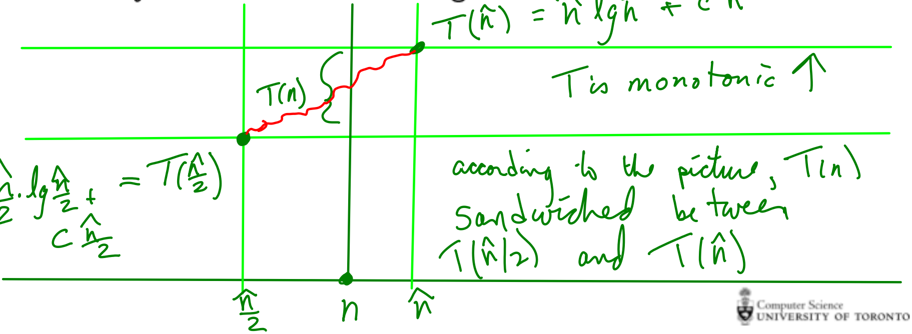
$n \in \mathbb{N},\ n \geq 2$

$\frac{\hat{n}}{2} < n \leq \hat{n} \longrightarrow \hat{n} = 2^{\lceil \lg n \rceil}$

next highest power of 2 for n

See Course Notes, Lemma 3.6 Exercise: Prove the recurrence for binary search is non-decreasing

$T(\hat{n}) = \hat{n} \lg \hat{n} + c\hat{n}$



T is monotonic ↑

$\frac{\hat{n}}{2} \cdot \lg \frac{\hat{n}}{2} + c\frac{\hat{n}}{2} = T(\frac{\hat{n}}{2})$

according to the picture, T(n) sandwiched between $T(\hat{n}/2)$ and $T(\hat{n})$

Prove $T \in O(n \lg n)$ for general case

$T(n) = T(\lceil n/2 \rceil) + T(\lfloor n/2 \rfloor) + n$ use $T(\hat{n}) = \hat{n} \lg \hat{n} + c\hat{n}$

pick $d = 2(2+c)$, then $d \in \mathbb{R}^+$

pick $B = 2$, then $B \in \mathbb{N}$

let $n \in \mathbb{N}$, $n \geq B$

See other bound in Notes

for Ω replace $\leq$ by $\geq$, $\hat{n}$, $\frac{\hat{n}}{2}$

$$
\begin{aligned}
T(n) &\leq T(\hat{n}) && \text{\# monotonicity of } T \\
&= \hat{n} \lg \hat{n} + c\hat{n} && \text{\# } \hat{n} \text{ is a power of 2} \\
& && \text{\# and students diligently} \\
&< 2n \lg(2n) + 2cn && \text{\# proved this.} \\
&= 2n(\lg(2) + \lg n) + 2nc && \text{\# } n > \tfrac{\hat{n}}{2} \Rightarrow 2n > \hat{n} \\
&= 2n((1+c) + \lg n) \\
&\leq 2n((1+c)\lg n + \lg n) && \text{\# } n \geq B \geq 2 \\
&= 2n(2+c) \lg n \\
&\leq d n \lg n && \text{\# } d \geq 2(2+c)
\end{aligned}
$$

$T(n) \in O(n \lg(n))$ Proving $T(n) \in \Omega(n \lg(n))$ similar because

$\hat{n} > n \Rightarrow \frac{\hat{n}}{2} > n/2$

What $\hat{n}$ means...

$\hat{n} = 2^{\lceil \lg n \rceil}$, i.e. next largest power of 2.

e.g. $n = 8$,

$\lg 8 = 3$

$\hat{n} = 2^{\lceil 3 \rceil} = 8$

$n = 7$

$\lg 7 = 2 < x < 3$

$\hat{n} = 2^{\lceil x \rceil} = 2^3 = 8$

$n = 1$, $\hat{n} = 1$

$n = 2$, $\hat{n} = 2$

$n = 3$, $\hat{n} = 4$

$n = 4$, $\hat{n} = 4$

$n = 5, 6, 7, 8$

$\hat{n} = 8$

$n = 9, 10, \ldots, 16$

$\hat{n} = 16$

General case

D+C algorithms
- fast mult
- big number
- big matrices
- closest pair of points.

Class of algorithms: partition problem into b *roughly* equal subproblems, solve, and recombine:

$$T(n) = \begin{cases} k & \text{if } n \leq B \\ a_1 T(\lceil n/b \rceil) + a_2 T(\lfloor n/b \rfloor) + f(n) & \text{if } n > B \end{cases}$$

where B, $k > 0$, $a_1, a_2 \geq 0$, and $a = a_1 + a_2 > 0$. $f(n)$ is the cost of splitting and recombining.

hope is polynomial.

Master Theorem

Binary search
$a = 1$
$b = 2$
$d = 0$
$a = b^d$

Merge Sort
$a = 2$
$b = 2$
$d = 1$
$a = b^d$

polynomial

$n^1 \cdot \log n$

base we divide problems by

If f from the previous slide has $f \in \theta(n^d)$, then

$$T(n) \in \begin{cases} \theta(n^d) & \text{if } a < b^d \\ \underline{\theta(n^d \log n)} & \text{if } a = b^d \\ \underline{\theta(n^{\log_b a})} & \boxed{\text{if } a > b^d} \end{cases}$$

$n^0 \lg n$
$= \lg n$

geometric
looks daunting.

Proof sketch

1. Unwind the recurrence, and prove a result for $n = b^k$
2. Prove that T is non-decreasing
3. Extend to all n, similar to MergeSort

Notes

Suppose we had proved that for $n = 2^k$ (power of 2) we know $T(n) = n\lg n + cn$. Want to show for all n (not just powers of 2), $T(n) \in \Omega(n \lg n)$

Pick $d = \underline{1/4}$. Then $d \in \mathbb{R}^+$ define $\hat{n} = 2^{\lceil \lg n \rceil}$

Pick $B = \underline{4}$. Then $B \in \mathbb{N}$. $\frac{\hat{n}}{2} < n \leq \hat{n}$

Let $n \in \mathbb{N}$. assume $n \geq B$.

$$T(n) \geq T\left(\frac{\hat{n}}{2}\right) \quad \# \text{ T monotonic},$$

$$= \frac{\hat{n}}{2} \cdot \lg \frac{\hat{n}}{2} + c\frac{\hat{n}}{2} \quad \# \text{ unwinding + induction}$$

$$\geq \frac{n}{2} \cdot \lg (n/2) + c\frac{n}{2} \quad \# \ n \leq \hat{n} \Rightarrow \frac{n}{2} < \frac{\hat{n}}{2}$$

$$= \frac{n}{2}(\lg n - \lg 2) + c\frac{n}{2} = n/2 \lg n - \frac{n}{2} + c\frac{n}{2}$$

$$\geq \frac{n}{4}\lg n + \frac{n}{4}\lg n - \frac{n}{2} + \frac{cn}{2} \quad \# \text{ split up } \frac{n}{2}\lg n$$

$$\geq n/4 \lg n + \frac{cn}{2} \quad \# \ n/4 \lg n \geq n/2 \quad n \geq B \geq 4$$

$$\geq n/4 \lg n \quad \# \ \frac{cn}{2} \geq 0$$