



Ayudantía 2

24 de agosto de 2018

Ayudante: Andrés Ríos

Teorema Maestro

Sea $f : \mathbb{N} \rightarrow \mathbb{R}_0^+$, $a, b, c \in \mathbb{R}_0^+$ tales que $a \geq 1$ y $b > 1$, y $T(n)$ una función definida por la siguiente ecuación de recurrencia:

$$T(n) = \begin{cases} c & n = 0 \\ a \cdot T(\lfloor \frac{n}{b} \rfloor) + f(n) & n \geq 1 \end{cases}$$

Se tiene que:

1. Si $f(n) \in O(n^{\log_b(a)-\varepsilon})$ para $\varepsilon > 0$, entonces $T(n) \in \Theta(n^{\log_b(a)})$
2. Si $f(n) \in \Theta(n^{\log_b(a)})$, entonces $T(n) \in \Theta(n^{\log_b(a)} \cdot \log_2(n))$
3. Si $f(n) \in \Omega(n^{\log_b(a)+\varepsilon})$ para $\varepsilon > 0$ y f es (a, b) -regular, entonces $T(n) \in \Theta(f(n))$.

f es (a, b) -regular si $\exists c \in \mathbb{R}^+, \exists n_0 \in \mathbb{N} : c < 1 \wedge (\forall n \geq n_0)(a \cdot f(\lfloor \frac{n}{b} \rfloor) \leq c \cdot f(n))$

Problema 1

Analice la complejidad de las siguientes recurrencias en notación O . El caso base para todas es $T(n) = 1$ cuando $n = 1$.

1. $T(n) = 2 \cdot T(\lfloor n/2 \rfloor) + n \log(n)$

Solución:

Para este caso no aplica el teorema maestro. Se hace por inducción constructiva.

En la ayudantía pasada vimos que cuando n es de la forma 2^i , entonces $T(n) \in O(n \log^2(n))$. Hagamos inducción para demostrar que $T(n) \leq c \cdot n \log^2(n)$ para algún c, n_0 .

Caso base $n = 1$:

$$T(1) = 1 \leq c \cdot 1 \cdot \log^2(1) = 0 \rightarrow \leftarrow$$

La propiedad no se cumple para $n = 1$. Intentemos $n = 2$:

$$T(2) = 2 + 2 \cdot \log(2) = 4 \leq c \cdot 2 \cdot \log^2(2) = c \cdot 2$$

Esto se cumple cuando $c \geq 2$.

Notar que $T(3)$ depende de $T(1)$, pero para $T(1)$ no se cumple la propiedad. Por lo tanto, $n = 3$ también es un caso base:

$$T(3) = 2 + 3 \log(3) \leq c \cdot 3 \cdot \log^2(3)$$

Esto se cumple cuando $c \geq \frac{2+3 \log(3)}{3 \log^2 3} \approx 1.43$

Ahora tenemos que hacer el paso inductivo:

$$\begin{aligned} T(n) &= 2 \cdot T(\lfloor n/2 \rfloor) + n \log(n) \\ &\leq 2c \lfloor n/2 \rfloor \log^2(\lfloor n/2 \rfloor) + n \log(n) \\ &\leq 2c \frac{n}{2} \log^2(n/2) + n \log(n) \\ &= cn \log^2(n/2) + n \log(n) \\ &= cn(\log(n) - 1)^2 + n \log(n) \\ &= cn \log^2(n) - 2cn \log(n) + cn + n \log(n) \end{aligned}$$

Necesitamos que $cn \log^2(n) - 2cn \log(n) + cn + n \log(n) \leq cn \log^2(n)$. Trabajemos esta inecuación:

$$\begin{aligned} cn \log^2(n) - 2cn \log(n) + cn + n \log(n) &\leq cn \log^2(n) \\ -2cn \log(n) + cn + n \log(n) &\leq 0 \\ c(-2n \log(n) + n) &\leq -n \log(n) \\ c &\geq \frac{n \log(n)}{2n \log(n) - n} \\ c &\geq \frac{\log(n)}{2 \log(n) - 1} \end{aligned}$$

Tenemos que $\frac{\log(n)}{2 \log(n) - 1}$ es decreciente, cuyo valor es 1 cuando $n = 2$. De acá se obtiene la restricción $c \geq 1$, lo que es más débil que la restricción anterior sobre c .

Así, con $c = 2, n_0 = 2$ se tiene que $\forall n \geq n_0 : T(n) \leq c \cdot n \log^2(n)$. Por lo tanto, $T(n) \in O(n \cdot \log^2(n))$

$$2. T(n) = T(\lfloor n/2 \rfloor) + T(\lceil n/2 \rceil) + 1$$

Solución:

La recurrencia mezcla piso y techo, por lo que el teorema maestro tampoco aplica directamente (según lo visto en clases). Si la recurrencia fuera $T'(n) = 2T(\lfloor n/2 \rfloor) + 1$, el teorema maestro sí aplicaría y la solución sería $O(n)$. Intentemos demostrar que $T(n) \leq cn$.

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

Necesitamos que $cn + 1 \leq cn$, pero esto obviamente no se cumple para ninguna elección de c . Intentemos hacer la hipótesis de inducción más fuerte restando un término de orden menor: $T(n) \leq cn - d$

$$\begin{aligned} T(n) &= T(\lfloor n/2 \rfloor) + T(\lceil n/2 \rceil) + 1 \\ &\leq c \lfloor n/2 \rfloor - d + c \lceil n/2 \rceil - d + 1 \\ &= cn - 2d + 1 \\ &\leq cn - d \end{aligned} \quad (d \geq 1)$$

Nos falta probar el caso base cuando $n = 1$:

$$T(1) = 1 \leq c - d$$

Esto se cumple siempre que $c \geq d + 1$

Así, tenemos que $\forall n \geq 1 : T(n) \leq 2n - 1$. Por lo tanto, $T(n) \in O(n)$.

3. $T(n) = 3T(\lfloor n/4 \rfloor) + n \log(n)$

En este caso se tiene que $\log_4(3) < 1$, por lo que $f(n) = n \log(n) \in \Omega(n^{\log_4(3)+\varepsilon})$ para $\varepsilon > 0$. Por lo tanto, estaríamos en el último caso del teorema maestro si f es $(3, 4)$ -regular.

Para demostrar que $f(n)$ es $(3, 4)$ -regular tenemos que encontrar $n_0 \in \mathbb{N}, c \in \mathbb{R}^+, c < 1$ tales que $\forall n \geq n_0$:

$$3f(\lfloor n/4 \rfloor) \leq cf(n)$$

Desarrollando se obtiene

$$\begin{aligned} 3f(\lfloor n/4 \rfloor) &= 3 \lfloor n/4 \rfloor \log(\lfloor n/4 \rfloor) \\ &\leq 3 \frac{n}{4} \log\left(\frac{n}{4}\right) \\ &= \frac{3}{4}n(\log(n) - 2) \\ &\leq \frac{3}{4}n \log(n) \end{aligned}$$

Así, con $n_0 = 1, c = \frac{3}{4}$ se tiene que $f(n) = n \log(n)$ es $(3, 4)$ -regular, por lo que utilizando el teorema maestro tenemos que $T(n) = \Theta(n \log(n))$.

Problema 2

Sea $T(n)$ la cantidad de operaciones realizadas por el algoritmo de Karatsuba al multiplicar dos números con n dígitos en el peor caso. Construya una ecuación de recurrencia para $T(n)$ y clasifíquela utilizando notación Θ .

No puede utilizar los supuestos y simplificaciones hechas en clases.

```
function KARATSUBA( $n_1, n_2, m$ )  
  if  $m = 1$  then  
    return  $n_1 \cdot n_2$   
   $m_2 \leftarrow \lfloor m/2 \rfloor$   
   $h_1, l_1 \leftarrow \text{split\_at}(n_1, m_2)$   
   $h_2, l_2 \leftarrow \text{split\_at}(n_2, m_2)$   
   $z_0 \leftarrow \text{Karatsuba}(l_1, l_2)$   
   $z_1 \leftarrow \text{Karatsuba}((l_1 + h_1), (l_2 + h_2))$   
   $z_2 \leftarrow \text{Karatsuba}(h_1, h_2)$   
  return  $z_2 \cdot 10^{2m_2} + (z_1 - z_2 - z_0) \cdot 10^{m_2} + z_0$ 
```

Solución:

Primero se demostrará que $T(n)$ es no decreciente haciendo inducción sobre n :

p.d. $\forall n \quad T(n) \leq T(n+1)$

- Casos Base: $n \leq 3 \rightarrow T(1) = 1 \leq T(2) = 2 \leq T(3) = 3 \leq T(4) = 2T(2) + T(3) + 4 = 11 \quad \checkmark$

- Hipotesis de Inducción: $\forall k \in \{1, \dots, n-1\} \quad T(k) \leq T(k+1)$
- Tesis de Inducción: $T(n) \leq T(n+1)$

$$\begin{aligned}
T(n) &= T\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + T\left(\left\lceil \frac{n}{2} \right\rceil\right) + T\left(\left\lceil \frac{n}{2} \right\rceil + 1\right) + n \\
&\leq T\left(\left\lfloor \frac{n+1}{2} \right\rfloor\right) + T\left(\left\lceil \frac{n+1}{2} \right\rceil\right) + T\left(\left\lceil \frac{n+1}{2} \right\rceil + 1\right) + n + 1 \\
&= T(n+1)
\end{aligned}$$

$$\implies T(n) \leq T(n+1)$$

Así, se concluye que $T(n)$ es no decreciente.

Utilizando esto, tenemos que

$$\begin{aligned}
T(n) &= T\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + T\left(\left\lceil \frac{n}{2} \right\rceil\right) + T\left(\left\lceil \frac{n}{2} \right\rceil + 1\right) + n \\
&\leq T\left(\left\lceil \frac{n}{2} \right\rceil + 1\right) + T\left(\left\lceil \frac{n}{2} \right\rceil + 1\right) + T\left(\left\lceil \frac{n}{2} \right\rceil + 1\right) + n \\
&= T\left(\left\lceil \frac{n}{2} + 1 \right\rceil\right) + T\left(\left\lceil \frac{n}{2} + 1 \right\rceil\right) + T\left(\left\lceil \frac{n}{2} + 1 \right\rceil\right) + n \\
&= 3 \cdot T\left(\left\lceil \frac{n}{2} + 1 \right\rceil\right) + n
\end{aligned}$$

$$\implies T(n) \leq 3 \cdot T\left(\left\lceil \frac{n}{2} + 1 \right\rceil\right) + n$$

Podemos hacer una transformación del dominio, definiendo una nueva recurrencia $S(n) = T(n+2)$.

Notar que $S(n)$ es también no decreciente. Tenemos que

$$\begin{aligned}
S(n) &= T(n+2) \leq 3 \cdot T\left(\left\lceil \frac{n+2}{2} + 1 \right\rceil\right) + n + 2 \\
&= 3 \cdot T\left(\left\lceil \frac{n}{2} + 2 \right\rceil\right) + n + 2 \\
&= 3 \cdot T\left(\left\lceil \frac{n}{2} \right\rceil + 2\right) + n + 2 \\
&= 3 \cdot S\left(\left\lceil \frac{n}{2} \right\rceil\right) + n + 2
\end{aligned}$$

$$\implies S(n) \leq 3 \cdot S\left(\left\lceil \frac{n}{2} \right\rceil\right) + n + 2$$

Así, podemos definir una tercera recurrencia:

$$R(n) = \begin{cases} S(10) & n \leq 10 \\ 3 \cdot R\left(\left\lceil \frac{n}{2} \right\rceil\right) + n + 2 & n > 10 \end{cases}$$

Proposición: $\forall n \geq 0 \quad R(n) \geq S(n)$

Demostración:

Hacemos inducción sobre n :

- Caso base: $n \leq 10$

$$R(n) = S(10) \geq S(n) \quad \checkmark$$

- Hipotesis de Inducción: $\forall k \in \{0, \dots, n-1\} \ R(k) \geq S(k)$
- Tesis de Inducción: $R(n) \geq S(n)$

$$\begin{aligned} R(n) &= 3 \cdot R\left(\left\lceil \frac{n}{2} \right\rceil\right) + n + 2 \\ &\geq 3 \cdot S\left(\left\lceil \frac{n}{2} \right\rceil\right) + n + 2 \\ &\geq S(n) \quad \blacksquare \end{aligned}$$

Como $f(n) = n + 2 \in O(n^{\log_2(3)-0.1})$, por el teorema maestro se deduce que $R(n) \in \Theta(n^{\log_2(3)})$. Como $S(n) \leq R(n) \ \forall n \geq 0$, entonces $S(n) \in O(n^{\log_2(3)})$.

Así, $T(n) = S(n-2) \in O((n-2)^{\log_2(3)}) = O(n^{\log_2(3)})$.

Por el otro lado, tenemos que

$$\begin{aligned} T(n) &= T\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + T\left(\left\lceil \frac{n}{2} \right\rceil\right) + T\left(\left\lceil \frac{n}{2} \right\rceil + 1\right) + n \\ &\geq T\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + T\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + T\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + n \\ &= 3 \cdot T\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + n \end{aligned}$$

Utilizando la misma idea anterior, podemos definir la recurrencia

$$U(n) = \begin{cases} T(0) & n \leq 10 \\ 3 \cdot U\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + n & n > 10 \end{cases}$$

Por inducción se deduce que $U(n) \leq T(n) \ \forall n \geq 0$:

- Caso base: $n \leq 10$

$$U(n) = T(0) \leq T(n) \quad \checkmark$$

- Hipotesis de Inducción: $\forall k \in \{0, \dots, n-1\} \ U(k) \leq T(k)$
- Tesis de Inducción: $U(n) \leq T(n)$

$$\begin{aligned} U(n) &= 3 \cdot U\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + n \\ &\leq 3 \cdot T\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + n \\ &\leq T(n) \quad \blacksquare \end{aligned}$$

Por el teorema maestro, $U(n) \in \Theta(n^{\log_2(3)})$. Como $U(n) \leq T(n) \ \forall n \geq 0$, tenemos que $T(n) \in \Omega(n^{\log_2(3)})$.

Así, $T(n) \in O(n^{\log_2(3)})$ y $T(n) \in \Omega(n^{\log_2(3)}) \implies T(n) \in \Theta(n^{\log_2(3)})$.

Problema 3

Considere la siguiente ecuación de recurrencia:

$$T(n) = \begin{cases} 1 & n = 1 \\ a \cdot T\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + b \cdot T\left(\left\lceil \frac{n}{2} \right\rceil\right) + c \cdot n^k & n > 1 \end{cases}$$

donde $a, b, k \in \mathbb{N}$, $a, b > 0$ y $c \in \mathbb{R}^+$.

Entregue una clasificación completa de $T(n)$ en notación Θ dependiendo de los valores de a , b y k .

Solución:

Primero se demuestra que $T(n)$ es no decreciente:

p.d. $\forall n \quad T(n) \leq T(n+1)$

- Caso base: $n = 1$

$$T(1) = 1 \leq T(2) = a + b + c$$

- Hipotesis de Inducción: $\forall k \in \{1, \dots, n-1\} \quad T(k) \leq T(k+1)$
- Tesis de Inducción: $T(n) \leq T(n+1)$

$$\begin{aligned} T(n) &= a \cdot T\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + b \cdot T\left(\left\lceil \frac{n}{2} \right\rceil\right) + c \cdot n^k \\ &\leq a \cdot T\left(\left\lfloor \frac{n+1}{2} \right\rfloor\right) + b \cdot T\left(\left\lceil \frac{n+1}{2} \right\rceil\right) + c \cdot (n+1)^k \\ &= T(n+1) \end{aligned}$$

Por lo tanto, $T(n)$ es no decreciente. Luego tenemos que

$$(a+b)T(\lfloor n/2 \rfloor) + cn^k \leq T(n) \leq (a+b)T(\lceil n/2 \rceil) + cn^k$$

Podemos definir las recurrencias auxiliares

$$\begin{aligned} T_1(n) &= \begin{cases} 1 & n = 1 \\ (a+b)T_1(\lfloor n/2 \rfloor) + cn^k & n > 1 \end{cases} \\ T_2(n) &= \begin{cases} 1 & n = 1 \\ (a+b)T_2(\lceil n/2 \rceil) + cn^k & n > 1 \end{cases} \end{aligned}$$

donde es trivial demostrar que $T_1(n) \leq T(n) \leq T_2(n)$. Tenemos que cuando el teorema maestro es aplicable, entonces la clasificación de $T_1(n)$ y $T_2(n)$ es idéntica. Si $T_1(n) \in \Theta(f(n))$ y $T_2(n) \in \Theta(f(n))$, entonces por la inecuación anterior también tendríamos que $T(n) \in \Theta(f(n))$. El único problema sería cuando el teorema maestro no aplica, lo que podría darse si $cn^k \in \Omega(n^{\log_2(a+b)+\varepsilon})$ pero cn^k no es $(a+b, 2)$ -regular. Sin embargo, esto no es posible:

$$(a+b)c \lfloor n/2 \rfloor^k \leq (a+b)c(n/2)^k = \frac{a+b}{2^k} cn^k$$

En este caso $k > \log_2(a+b)$, por lo que $\frac{a+b}{2^k} < 1$ sería una constante válida para la definición de regularidad. En otras palabras, si se cae en el tercer caso del teorema maestro, entonces cn^k siempre es $(a+b, 2)$ -regular.

Con esto, la clasificación queda dada por el teorema maestro:

- Si $k < \log_2(a+b)$, entonces $T(n) \in \Theta(n^{\log_2(a+b)})$
- Si $k = \log_2(a+b)$, entonces $T(n) \in \Theta(n^{\log_2(a+b)} \log_2(n))$
- Si $k > \log_2(a+b)$, entonces $T(n) \in \Theta(n^k)$

Problema 4

Se busca analizar la complejidad de BucketSort. El arreglo A de entrada contiene números entre el 0 y el 1.

```
function BUCKETSORT( $A[0\dots n-1]$ )  
  for  $i = 0\dots n$  do  
     $B[i] \leftarrow \emptyset$   
  for  $i = 0\dots n$  do  
    Insertar  $A[i]$  en la lista  $B[\lfloor A[i] \cdot n \rfloor]$   
  for  $i = 0\dots n$  do  
    Ordenar  $B[i]$  con InsertionSort  
  return Concatenación de  $B$  en orden
```

Demuestre que si los valores de A distribuyen uniforme entre 0 y 1, entonces BucketSort funciona en $\Theta(n)$ en promedio.

Idea de solución:

Sea n_i una variable aleatoria que indica la cantidad de elementos en el bucket i . Como *InsertionSort* funciona en tiempo $O(n^2)$, entonces el tiempo de *BucketSort* está dado por (relajando la notación):

$$T(n) = \Theta(n) + \sum_{i=0}^{n-1} O(n_i^2)$$

Luego, el tiempo promedio del algoritmo es

$$\begin{aligned}\mathbb{E}[T(n)] &= \mathbb{E}\left[\Theta(n) + \sum_{i=0}^{n-1} O(n_i^2)\right] \\ &= \Theta(n) + \sum_{i=0}^{n-1} \mathbb{E}[O(n_i^2)] \\ &= \Theta(n) + \sum_{i=0}^{n-1} O(\mathbb{E}[n_i^2])\end{aligned}$$

Necesitamos resolver $\mathbb{E}[n_i^2]$. Para esto se define el conjunto de variables aleatorias X_{ij} , donde

$$X_{ij} = \begin{cases} 1 & A[j] \text{ se asignó al bucket } B[i] \\ 0 & \text{en otro caso} \end{cases}$$

Se tiene que

$$\begin{aligned}\mathbb{E}[n_i^2] &= \mathbb{E}\left[\left(\sum_{j=0}^{n-1} X_{ij}\right)^2\right] \\ &= \mathbb{E}\left[\sum_{j=0}^{n-1} X_{ij}^2 + \sum_{j=0}^{n-1} \sum_{\substack{k=0 \\ k \neq j}}^{n-1} X_{ij} X_{ik}\right] \\ &= \sum_{j=0}^{n-1} \mathbb{E}[X_{ij}^2] + \sum_{j=0}^{n-1} \sum_{\substack{k=0 \\ k \neq j}}^{n-1} \mathbb{E}[X_{ij} X_{ik}]\end{aligned}$$

Luego,

$$\mathbb{E}[X_{ij}^2] = 1^2 \cdot \Pr[X_{ij} = 1] + 0^2 \cdot \Pr[X_{ij} = 0] = \frac{1}{n}$$

Por el otro lado, si $j \neq k$ entonces X_{ij} es independiente de X_{ik} , por lo que

$$\mathbb{E}[X_{ij}X_{ik}] = \mathbb{E}[X_{ij}]\mathbb{E}[X_{ik}] = \frac{1}{n^2}$$

Reemplazamos:

$$\begin{aligned}\mathbb{E}[n_i^2] &= \sum_{j=0}^{n-1} \mathbb{E}[X_{ij}^2] + \sum_{j=0}^{n-1} \sum_{\substack{k=0 \\ k \neq j}}^{n-1} \mathbb{E}[X_{ij}X_{ik}] \\ &= \sum_{j=0}^{n-1} \frac{1}{n} + \sum_{j=0}^{n-1} \sum_{\substack{k=0 \\ k \neq j}}^{n-1} \frac{1}{n^2} \\ &= \frac{1}{n} \sum_{j=0}^{n-1} 1 + \frac{1}{n^2} \sum_{j=0}^{n-1} \sum_{\substack{k=0 \\ k \neq j}}^{n-1} 1 \\ &= \frac{n}{n} + \frac{n(n-1)}{n^2} \\ &= 2 - \frac{1}{n}\end{aligned}$$

Por lo tanto,

$$\mathbb{E}[T(n)] = \Theta(n) + O(n(2 - 1/n)) \in \Theta(n)$$