Analyse de la complexité du tri fusion

Le principe du tri fusion consiste à trier récursivement la moitié gauche d'un tableau, puis la moitié droite, puis d'interclasser les deux moitiés du tableau. L'interclassement a une complexité $O(n)$ où n est la taille du tableau résultat.

L'équation de récurrence de la complexité dans le pire des cas est donc

$$C_{fusion}(n) = \alpha n + \beta + C_{fusion}(\left\lceil \frac{n}{2} \right\rceil) + C_{fusion}(\left\lfloor \frac{n}{2} \right\rfloor)$$

on cherche à montrer que:

$$\exists a, \quad C_{fusion}(n) \leq an(\log_2(n) + 1) \tag{1}$$

Lorsque $n = 1$, il suffit de choisir la constante a de sorte que $a \geq C_{fusion}(1)$

Soit $k \geq 2$. Supposons que la propriété 1 soit vraie pour toutes les valeurs de $n \leq k$.

Soit $n = k + 1$

1 Cas où n est pair

Nous avons

$$C_{fusion}(n) = \alpha n + \beta + 2C_{fusion}(\frac{n}{2})$$

En utilisant l'hypothèse de récurrence, comme $1 \leq \frac{n}{2} \leq k$, on peut donc dire que

$$C_{fusion}(n) \leq \alpha n + \beta + 2a\frac{n}{2}(\log_2(\frac{n}{2}) + 1)$$

Comme $\log_2(\frac{n}{2}) = \log_2(n) - 1$, on a donc

$$C_{fusion}(n) \leq \alpha n + \beta + an(\log_2(n))$$

En choisissant a de sorte que $a \geq \alpha + \beta$, on a $\alpha n + \beta \leq an$ et donc

$$C_{fusion}(n) \leq an(\log_2(n) + 1)$$

On remarque en particulier que cette propriété est donc vraie pour $n = 2$

2 Cas où n est impair

Dans ce cas, l'équation de récurrence devient:

$$C_{fusion}(n) = \alpha n + \beta + C_{fusion}(\frac{n+1}{2}) + C_{fusion}(\frac{n-1}{2})$$

Comme $2 \leq \frac{n+1}{2} \leq k$ et $1 \leq \frac{n-1}{2} \leq k$, on peut appliquer l'hypothèse de récurrence.

D'où:

$$C_{fusion}(n) \leq \alpha n + \beta + a\frac{n+1}{2}(\log_2(\frac{n+1}{2}) + 1) + a\frac{n-1}{2}(\log_2(\frac{n-1}{2}) + 1)$$

En développant les logarithmes de cette expression, on trouve:

$$C_{fusion}(n) \leq \alpha n + \beta + a\frac{n+1}{2}(\log_2(n+1)) + a\frac{n-1}{2}(\log_2(n-1))$$

Le logarithme est une fonction croissante. On peut donc considérer que $\log_2(n+1) = \log_2(n) + \epsilon$, et $\log_2(n-1) = \log_2(n) - \epsilon'$.

De plus, comme la pente du logarithme est une fonction décroissante, pour $n \geq 3$, $\epsilon \leq \epsilon' \leq \log_2(3) - \log_2(2) = b < 1$

D'où l'expression:

$$C_{fusion}(n) \leq \alpha n + \beta + a\frac{n+1}{2}(\log_2(n) + \epsilon) + a\frac{n-1}{2}(\log_2(n) - \epsilon')$$

Et donc

$$C_{fusion}(n) \leq \alpha n + \beta + an\log_2(n) + a(\epsilon\frac{n+1}{2} - \epsilon'\frac{n-1}{2})$$

On peut donc écrire

$$C_{fusion}(n) \leq \alpha n + \beta + an\log_2(n) + a\epsilon'(\frac{n+1}{2} - \frac{n-1}{2}) \leq \alpha n + \beta + an\log_2(n) + ab$$

Considérons $\alpha n + \beta + ab \leq (\alpha + \beta)n + ab$. On souhaite que $a(n-b) \geq (\alpha+\beta)n$ pour tout $n \geq 3$. Pour cela il suffit de considérer que $\frac{n}{n-b} = 1 + \frac{b}{n-b}$ voit sa valeur maximale pour la valeur minimale de n. Or on a supposé que $n \geq 3$.

En choisissant donc $a \geq (\alpha+\beta)\frac{3}{3-b}$ (notez que cela ne dépend pas de n), on obtient $alphan + \beta + ab \leq an$ pour tout n. D'où

$$C_{fusion}(n) \leq an\log_2(n) + an = an(\log_2(n) + 1)$$