DIVISER POUR RÉGNER

À la fin de ce chapitre, je sais :

- ☞ énoncer le principe d'un algorithme de type diviser pour régner
- ☞ distinguer les cas d'utilisation de ce principe (sous-problèmes indépendants)
- ☞ évaluer la complexité d'un algorithme diviser pour régner

A Diviser pour régner

■ **Définition 1 — Algorithme de type diviser pour régner.** L'idée centrale d'un algorithme de type diviser pour régner est de décomposer le problème étudié en plusieurs sous-problèmes de taille réduite. Ces algorithmes peuvent éventuellement être exprimés récursivement et sont souvent très efficaces.

On distingue trois étapes lors de l'exécution d'un tel algorithme :

1. la division du problème en sous-problèmes qu'on espère plus simples à résoudre,
2. la résolution des sous-problèmes, c'est à cette étape que l'on peut faire appel à la récursivité,
3. la combinaison des solutions des sous-problèmes pour construire la solution au problème.

Ⓡ On devrait donc logiquement nommer ces algorithmes diviser, résoudre et combiner !

Vocabulary 1 — Divide and conqueer ⇜⇝ diviser pour régner.

Très souvent[1], les algorithmes de type diviser pour régner sont exprimés récursivement. À partir d'une taille de problème $\mathcal{P}$ de taille n, la division en sous-problèmes aboutit soit à $n = 1$ soit à $n = s$, s étant alors un taille pour laquelle on sait résoudre le problème facilement et efficacement. Les étapes 7 et 9 de l'algorithme 1 sont facilement descriptibles à l'aide d'un arbre, comme le montre la figure 2. La hauteur de cet arbre peut être appréciée et quantifiée. Elle sert notamment à calculer la complexité liée aux algorithme récursifs.

1. mais pas toujours

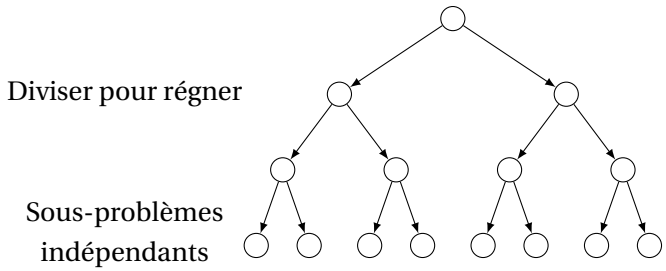

FIGURE 1 – Principe de la décomposition d'un problème en sous-problèmes indépendants pour un algorithme de type diviser pour régner.

Algorithme 1 Diviser, résoudre et combiner (Divide And Conquer)

```
1 : Fonction DRC(𝒫)                                   ▷ 𝒫 est un problème de taille n
2 :     r ← un entier ⩾ 1                  ▷ pour générer r sous-problèmes à chaque étape
3 :     d ← un entier > 1                              ▷ on divise par d la taille du problème
4 :     si n < s alors                    ▷ Condition d'arrêt, s est un seuil à déterminer
5 :         renvoyer RÉSOUDRE(n)
6 :     sinon
7 :         (𝒫_1, ..., 𝒫_r) ← Diviser 𝒫 en r sous problèmes de taille n/d.
8 :         pour i de 1 à r répéter
9 :             S_i ← DRC(𝒫_i)                                        ▷ Appels récursifs
10 :        renvoyer COMBINER(𝒮_1, ..., 𝒮_r)
```

Sur la figure 2, on a choisi de représenter un algorithme de type diviser pour régner dont le problème associé $\mathcal{P}$ est de taille n. L'étape de division en sous-problèmes divise par d le problème initial et nécessite r appels récursifs. Si $D(n)$ est la complexité de la partie division et $C(n)$ la complexité de l'étape de combinaison des résultats, alors on peut décrire la complexité $T(n)$ d'un tel algorithme par la relation de récurrence $T(n) = rT(n/d) + D(n) + C(n)$ et $T(s)$ une constante. Sur la figure 2 on a choisi $r = 3$ pour la représentation graphique.

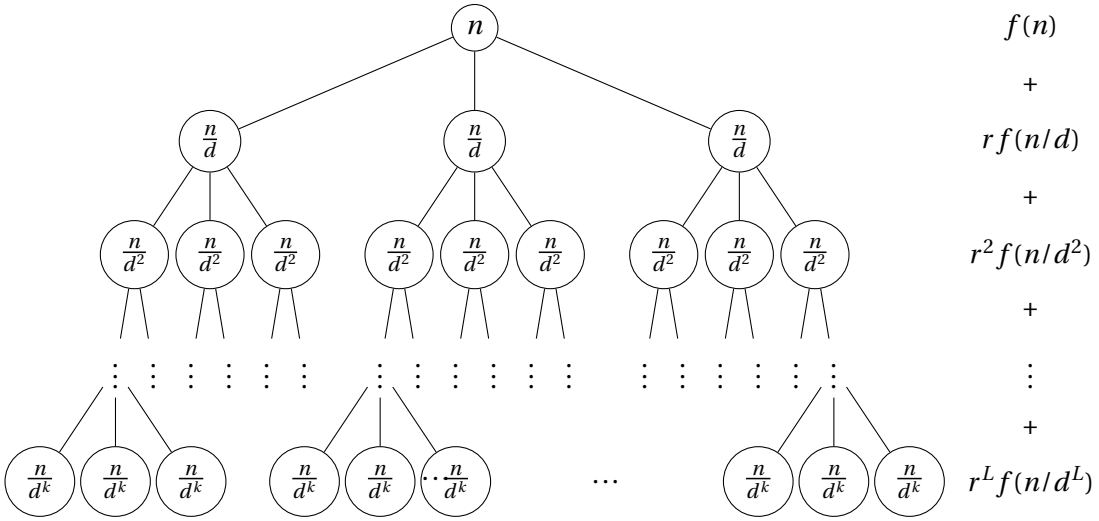

FIGURE 2 – Structure d'arbre et appels récursifs pour la récurrence : $T(n) = rT(n/d) + f(n)$ ET $n/d^k = s$. On a choisi $r = 3$ pour l'illustration, c'est à dire chaque nœud possède trois enfants au maximum : on opère trois appels récursifs à chaque étape de l'algorithme.

B Exemple de la recherche dichotomique

■ **Exemple 1 — Recherche dichotomique.** L'algorithme de recherche dichotomique 2 est un exemple d'algorithme de type diviser pour régner : la division du problème en sous-problèmes est opérée via la ligne 8. La résolution des sous-problèmes est effectuée par des appels récursifs. La combinaison des résultats n'est pas explicite mais s'effectue sur le tableau lui-même grâce aux indices g et d.

R La recherche dichotomique est donc bien un cas particulier d'algorithme diviser pour régner avec $r = 1$ et $d = 2$, c'est à dire un seul appel récursif par chemin d'exécution et une division de la taille du problème par deux.

Algorithme 2 Recherche récursive d'un élément par dichotomie dans un tableau trié

```
1 : Fonction REC_DICH(t, g, d, elem)
2 :     si g = d alors                                  ▷ Condition d'arrêt
3 :         si t[g] = elem alors
4 :             renvoyer g
5 :         sinon
6 :             renvoyer l'élément n'a pas été trouvé
7 :     sinon
8 :         m ← (g+d)//2                                ▷ Diviser
9 :         si t[m] < elem alors
10 :            REC_DICH(t, m+1, d, elem)               ▷ résoudre
11 :        sinon
12 :            REC_DICH(t, g, m, elem)                 ▷ résoudre
```

Grâce à la figure 3, on peut calculer le nombre d'opérations élémentaires $T(n)$ nécessaires à l'exécution de l'algorithme 2. On fait l'hypothèse que, hors appel récursif, la fonction REC_DICH nécessite un nombre constant d'opérations c. On peut expliciter formellement la relation de récurrence qui existe entre $T(n)$ et $T(n/2)$: on a $T(n) = T(n/2) + c$. on peut donc écrire :

$$
\begin{aligned}
T(n) &= T(n/2) + c && (1)\\
&= T(n/4) + c + c = T(n/4) + 2c && (2)\\
&= T(n/8) + 3c && (3)\\
&= \dots && (4)\\
&= T(n/2^k) + kc && (5)\\
\text{ă} &= T(1) + kc && (6)
\end{aligned}
$$

D'après l'algorithme 2, la condition d'arrêt s'effectue en un nombre constant d'opérations : $T(1) = O(1)$. Donc on a $T(n) = O(k)$. Or, on a $\frac{n}{2^k} = 1$. Donc $k = \log_2 n$ et $T(n) = O(\log n)$.

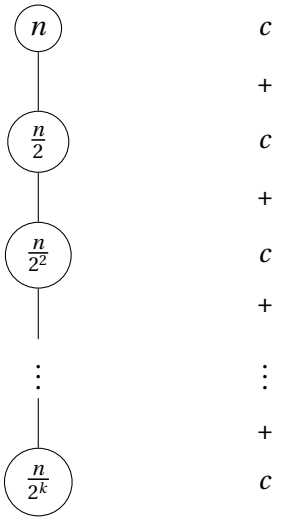

FIGURE 3 – Structure d'arbre et appels récursifs pour la récurrence de la recherche dichomotique : $T(n) = T(n/2) + c$ et $\frac{n}{2^k} = 1$. Hors appel récursif, la fonction opère un nombre constant d'opérations c.

C Exemple de l'exponentiation rapide

L'algorithme naïf de l'exponentiation (cf. algorithme 3) qui permet d'obtenir a^n en multipliant a par lui-même n fois n'est pas très efficace : sa complexité étant en $O(n)$.

Algorithme 3 Exponentiation naïve a^n

```
1 : Fonction EXP_NAIVE(a,n)
2 :     api ← 1
3 :     pour i de 0 à n − 1 répéter
4 :         api ← api × a
5 :     renvoyer api
```

Or, l'exponentiation est une opération très récurrente qu'il est nécessaire de pouvoir exécuter le plus rapidement possible. L'exponentiation rapide (cf. algorithme 4) propose une version récursive de type diviser pour régner dont la complexité est en $O(\log n)$.

L'analyse de l'algorithme 4 montre que :

- c'est un cas particulier d'algorithme diviser pour régner avec $r = 1$ et $d = 2$, c'est à dire un seul appel récursif par chemin d'exécution et une division de la taille du problème par deux[2],
- l'évolution du coût ne dépend pas de a mais de n, c'est à dire l'exposant.

On peut procéder de la même manière qu'avec l'algorithme 2 pour calculer la complexité et s'appuyer sur l'arbre de la figure 3. Pour simplifier le calcul, on peut considérer que la taille du

2. à un près si n est pair

Algorithme 4 Exponentiation rapide a^n

```
1 : Fonction EXP_RAPIDE(a,n)
2 :     si n = 0 alors                                  ▷ Condition d'arrêt
3 :         renvoyer 1
4 :     sinon si n est pair alors
5 :         p ← EXP_RAPIDE(a, n//2)                     ▷ Appel récursif
6 :         renvoyer p × p
7 :     sinon
8 :         p ← EXP_RAPIDE(a, (n-1)//2)                 ▷ Appel récursif
9 :         renvoyer p × p × a
```

problème est divisée par deux. Le coût hors appel récursif est constant car il s'agit de multiplications. On a donc $T(n) = O(\log n)$.

R L'algorithme 4 n'est pas à récursivité terminale. Sa version itérative est moins évidente. Celle-ci est détaillée par l'algorithme 5

Algorithme 5 Exponentiation rapide a^n (version itérative)

```
1 : Fonction EXP_RAPIDE_ITE(a, n)                       ▷ a et n sont des entiers naturels
2 :     si n = 0 alors
3 :         renvoyer 1
4 :     p ← a
5 :     m ← 1
6 :     n_it ← ⌊log_2(n)⌋
7 :     pour i de 1 à n_it répéter
8 :         m ← n//2^(n_it − i)
9 :         si m est pair alors
10 :            p ← p × p
11 :        sinon
12 :            p ← p × p × a
13 :    renvoyer p
```

D Exemple du tri fusion

Les tris génériques abordés jusqu'à présent, par sélection ou insertion, présentent des complexités polynomiales en $O(n^2)$ dans le pire des cas. L'algorithme de tri fusion a été inventé par John von Neumann en 1945. C'est un bel exemple d'algorithme de type diviser pour régner avec $r = 2$ et $d = 2$, c'est à dire deux appels récursifs par chemin d'exécution et une division de la taille du problème par deux (cf . figure 4). Il permet de dépasser cette limite et d'obtenir un tri générique de complexité logarithmique. Ce tri est comparatif, il peut s'effectuer en place et

les implémentations peuvent être stables.

Son principe (cf. algorithmes 6 et 7) est simple : transformer le tri d'un tableau à n éléments en sous-tableaux ne comportant qu'un seul élément [3] puis les recombiner en un seul tableau en conservant l'ordre. L'algorithme est divisé en deux fonctions :

- TRI_FUSION qui opère concrètement la division et la résolution des sous-problèmes,
- FUSION qui combine les solutions des sous-problèmes en fusionnant deux sous-tableaux triés.

Pour le calcul de la complexité, on a la relation de récurrence $T(n) = 2T(n/2) + f(n)$ où $f(n)$ représente le nombre d'opérations élémentaires nécessaires pour fusionner deux sous-tableaux de taille $n/2$. La condition d'arrêt de la récursivité est atteinte pour $\frac{n}{2^k} = 1$ c'est à dire pour $k = \log_2 n$. Par ailleurs, on remarque que $T(1) = c$ est un opération à coût constant, suite à la condition d'arrêt on ne fait que retourner le tableau non modifié.

$$
\begin{aligned}
T(n) &= 2T(n/2) + f(n) && (7)\\
&= 4T(n/4) + 2O(n) && (8)\\
&= 8T(n/8) + 8O(n) && (9)\\
&= \dots && (10)\\
&= 2^k T(1) + 2^k O(n) && (11)\\
&= c\log_2 n + O(n)\log_2 n && (12)
\end{aligned}
$$

Dans le pire des cas, l'étape de combinaison effectue la combinaison de deux sous-tableaux de dimension $n/2$ et les trois boucles de la fonction FUSION effectue $n/2$ tours. On a alors $f(n) = O(n)$. C'est pourquoi, la complexité du tri fusion vaut $T(n) = O(\log_2 n + n\log_2 n) = O(n\log n)$

Algorithme 6 Tri fusion

```
1 : Fonction TRI_FUSION(t, g, d)
2 :     si d = g alors                          ▷ Condition d'arrêt
3 :         renvoyer t
4 :     m ← (g+d)//2                            ▷ on découpe au milieu
5 :     TRI_FUSION(t, g, m)
6 :     TRI_FUSION(t, m+1, d)
7 :     FUSION(t, g, m, sup)
```

3. et donc déjà triés !

Algorithme 7 Fusion de sous-tableaux triés

```
1 : Fonction FUSION(t, g, m, d)
2 :     i ← m - g +1
3 :     j ← d - m
4 :     G, D deux tableaux de taille i et j
5 :     pour k de 1 à i répéter
6 :         G[k] ← t[g + k -1]
7 :     pour k de 1 à i répéter
8 :         D[k] ← t[m + k ]
9 :     G[i +1] ← None
10 :    D[j +1] ← None
11 :    p ← 1
12 :    q ← 1
13 :    pour v de g à d répéter
14 :        si G[p] < D[q] alors
15 :            t[v] ← G[p]
16 :            p ← p + 1
17 :        sinon
18 :            t[v] ← D[q]
19 :            q ← q + 1
```

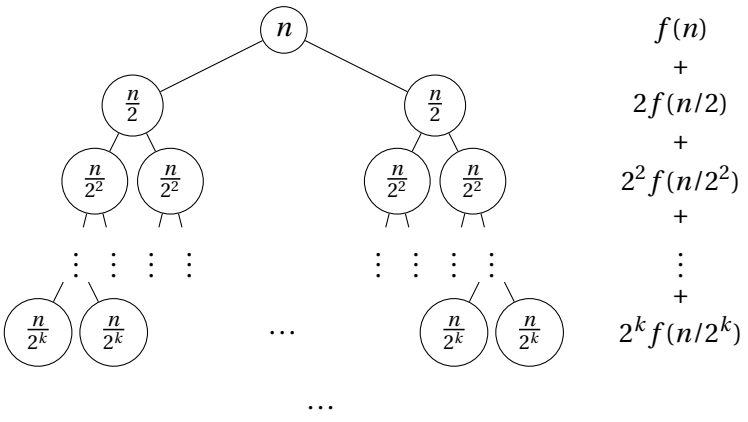

FIGURE 4 – Structure d'arbre et appels récursifs pour le tri fusion : $T(n) = 2T(n/2) + f(n)$ et $\frac{n}{2^k} = 1$. La fonction FUSION opère un nombre d'opérations $f(n)$.