

3I-IN9 : Structures de données

TD - 3

1 Recherche d'une valeur dans un tableau non-trié

Algorithm 1 Recherche dans un tableau non-trié

```

1: procedure FINDVALUE( $T[], v$ )
2:    $n \leftarrow \text{getNbElement}(T)$ 
3:   for  $i \leftarrow 1, n$  do
4:     if  $T[i] == v$  then
5:       return  $i$ 
6:     end if
7:   end for
8:   return  $-1$ 
9: end procedure

```

Dans cet exercice nous étudierons l'algorithme 1 qui permet de rechercher une valeur dans un tableau non-trié.

1.1. Analyse préliminaire de l'algorithme :

- 1.1.1. Pour un tableau de taille n , quelle est le critère sur le contenu du tableau qui fait varier le coût en temps de cet algorithme ?
- 1.1.2. La boucle à la ligne 3, est-elle une boucle bornée ou non-bornée ?
- 1.1.3. Le nombre d'itération est toujours égale à n (la taille du tableau) ?

1.2. Détermination du meilleur cas :

- 1.2.1. quelle est la condition pour que le meilleur cas se produise ?
- 1.2.2. Quel est son coût en temps ?
- 1.2.3. Quelle est sa complexité majorante asymptotique avec la notation de Landau ?

1.3. Détermination du pire cas :

- 1.3.1. Quelle est la condition pour que le pire cas se produise ?
- 1.3.2. Quel est son coût en temps ?
- 1.3.3. Quelle est sa complexité majorante asymptotique avec la notation de Landau ?

1.4. Détermination du cas moyen :

Rappel

Pour obtenir le coût moyen d'un algorithme qui traite un tableau de taille n , il faut calculer la moyenne des coûts pour tous les tableaux de taille n possible pondéré par la probabilité d'apparition d'un tel tableau. Cela s'écrit :

$$T_{\text{moy.}}(n) = \sum_{d \in \mathcal{D}_n} p(d) \text{Coût}(d), \quad (1)$$

avec \mathcal{D}_n l'ensemble des données de taille n (dans notre cas l'ensemble des tableaux de taille n), $p(d)$ la probabilité que le tableau d existe et $\text{Coût}(d)$ le coût de traitement du tableau d par l'algorithme.

Réduction du nombre de cas à traiter

Très rapidement, si nous tentons d'appliquer directement l'équation 1, nous remarquons que le nombre de cas à traiter est excessif. Par exemple, si nous considérons l'ensemble des tableaux contenant n entiers de 8 bits, pour $n = 10$ nous avons 1208925819614629174706176 cas à traiter. Il faut alors trouver **comment réduit le nombre de cas à traiter**, pour cela nous avons deux possibilités :

- faire des hypothèses "réalistes" sur l'ensemble des données pour réduire sa taille ;
- trouver des familles de tableaux qui ont le même coût.

- 1.4.1. Quelle(s) hypothèse(s) pouvons-nous faire pour réduire le nombre de cas à traiter ?
- 1.4.2. À partir de votre réponse à la question 1.1.1, comment pouvons-nous regrouper les tableaux pour diminuer le nombre de cas à traiter ?
- 1.4.3. Quelle hypothèse pouvez vous faire sur la probabilité d'apparition de vos groupes de tableaux ?
- 1.4.4. Calculez le coût en temps pour vos groupes de tableaux ?
- 1.4.5. Calculez le coût en temps moyen de l'algorithme ?
- 1.4.6. Quelle est sa complexité majorante asymptotique avec la notation de Landau ?

2 Recherche d'une valeur dans un tableau trié

Algorithm 2 Recherche par dichotomie

```
1: procedure FINDVALUE( $T[], v$ )
2:    $start \leftarrow 1$ 
3:    $end \leftarrow \text{getNbElement}(T)$ 
4:   while  $start \leq end$  do
5:      $i \leftarrow \lfloor \frac{start+end}{2} \rfloor$ 
6:     if  $T[i] == v$  then
7:       return  $i$ 
8:     else if  $T[i] < v$  then
9:        $start \leftarrow i + 1$ 
10:    else
11:       $end \leftarrow i - 1$ 
12:    end if
13:  end while
14:  return  $-1$ 
15: end procedure
```

2.1. Analyse préliminaire de l'algorithme :

- 2.1.1. Pour un tableau de taille n , quelle est le critère sur le contenu du tableau qui fait varier le coût en temps de cet algorithme ?
- 2.1.2. La boucle à la ligne 4, est-elle une boucle bornée ou non-bornée ?
- 2.1.3. À chaque itération de la boucle ligne 4 le tableau est découpé. Soit n_k la taille du tableau à la k -ème itération, donnez l'équation de la taille du sous-tableau à la $k + 1$ -ème itération ?
- 2.1.4. À partir de l'équation de la question 2.1.3, quelle hypothèse simplificatrice peut-on faire sur la taille initiale du tableau ?
- 2.1.5. Quelle est alors le nombre maximum d'itération que peut faire la boucle ligne 4 ?

2.2. Détermination du meilleur cas :

- 2.2.1. quelle est la condition pour que le meilleur cas se produise ?
- 2.2.2. Quel est son coût en temps ?
- 2.2.3. Quelle est sa complexité majorante asymptotique avec la notation de Landau ?

2.3. Détermination du pire cas :

2.3.1. Quelle est la condition pour que le pire cas se produise ?

2.3.2. Quel est son coût en temps ?

2.3.3. Quelle est sa complexité majorante asymptotique avec la notation de Landau ?

2.4. Détermination du cas moyen :

2.4.1. Quelle(s) hypothèse(s) pouvons-nous faire pour réduire le nombre de cas à traiter ?

2.4.2. À partir de votre réponse à la question 2.1.1, comment pouvons-nous regrouper les tableaux pour diminuer le nombre de cas à traiter ?

2.4.3. Quelle hypothèse pouvez vous faire sur la probabilité d'apparition de vos groupes de tableaux ?

2.4.4. Calculez le coût en temps pour vos groupes de tableaux ?

2.4.5. Calculez le coût en temps moyen de l'algorithme ?

2.4.6. Quelle est sa complexité majorante asymptotique avec la notation de Landau ?