

Information, Calcul et Communication

CS-119(k) ICC – Théorie Semaine 2

Rafael Pires
rafael.pires@epfl.ch

Précédemment, dans ICC-T 1



- Les ingrédients de base des algorithmes
 - **Données** (variables)
 - **Instructions** (affectations, structures de contrôle)
- Problèmes :
 - Calcul du **modulo 3** d'un grand nombre
 - Recherche du **minimum dans une liste**
 - Problème du **voyageur du commerce**
 - Comparaison **tous contre tous**
 - L'algorithme d'Euclide (**pgdc**)

Précédemment, dans ICC-T 1



- Les ingrédients de base des algorithmes
 - **Données** (variables)
 - **Instructions** (affectations, structures de contrôle)

Branchements



si $\text{delta} < 0$, alors ...
sinon ...

Boucles



pour i allant de 1 à n , (répéter ...)

tant que $i \leq 10$, (répéter ...)

Précédemment, dans ICC-T 1



Boucles

Itérations

pour i allant de 1 à 10,
(répéter ...)

itère **sur une**
séquence définie à
l'avance

Boucles conditionnelles

$i \leftarrow 1$
tant que $i \leq 10$
(répéter ...)
 $i \leftarrow i + 1$

condition peut être
n'importe quelle
expression booléenne

Aujourd'hui

- Sous-algorithmes
- Complexité temporelle
- Notation Grand Theta

Sous-algorithmes



Sous-algorithmes Sous-recettes

- Préparer le pain → Peut être une **sous-recette** (faire du pain maison)
- Couper les légumes → Une autre **sous-recette**
(utilisable pour une salade aussi)
- Assembler le sandwich → Une tâche qui **utilise les résultats** des
sous-recettes précédentes

Sous-algorithmes

Algorithme principal

entrée: ...
sortie: ...

...
 $x \leftarrow \text{algo1}(5)$
 $y \leftarrow 10 + \text{algo2}(7)$
 $x \leftarrow x + \text{algo1}(y)$
...

algo1

entrée: nombre entier n
sortie: nombre entier

...

algo2

entrée: nombre entier n
sortie: nombre entier

...

Illustration du principe avec le tri d'une liste

- Comment trier une liste de nombres ?
- Il existe de nombreuses façons de faire, plus ou moins efficaces. Nous allons en voir une : **le tri par insertion**, qui permet de bien illustrer le principe de l'utilisation de sous-algorithmes.

Illustration du principe avec le tri d'une liste

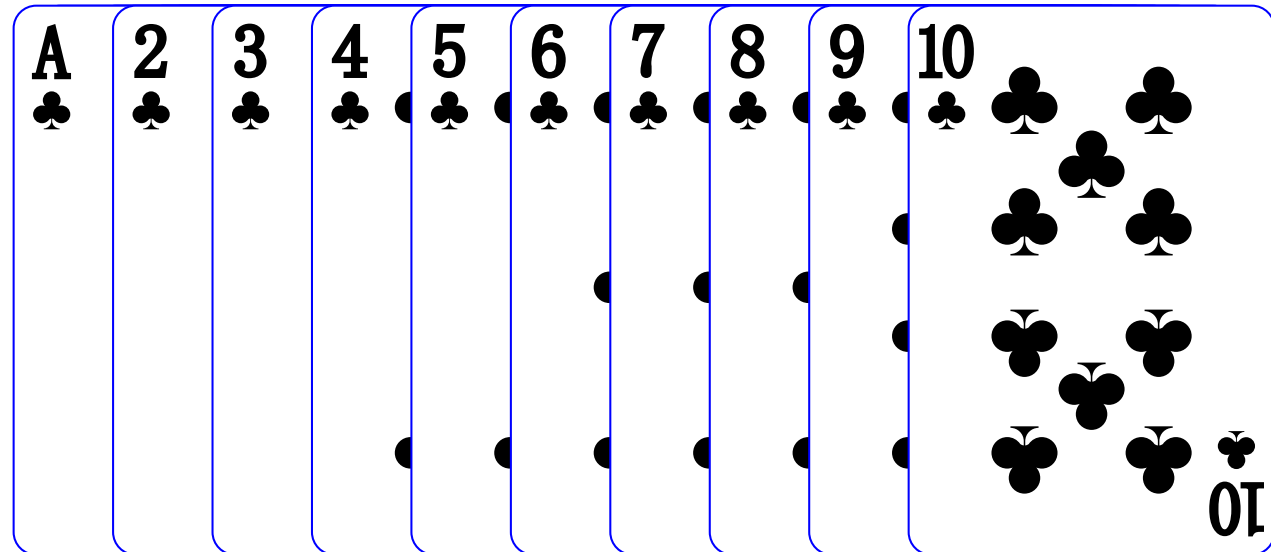
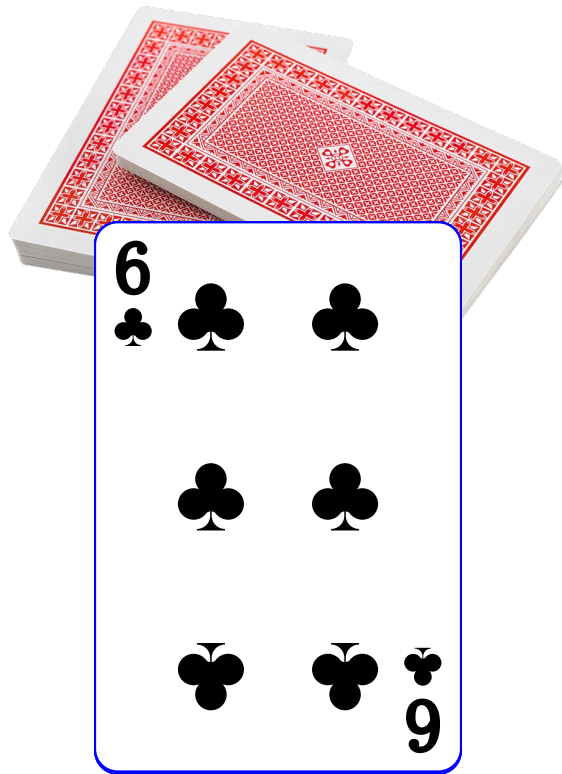


Illustration du principe avec le tri d'une liste



Tri d'une liste – 1^{er} essai

entrée : liste **L** de taille **n**

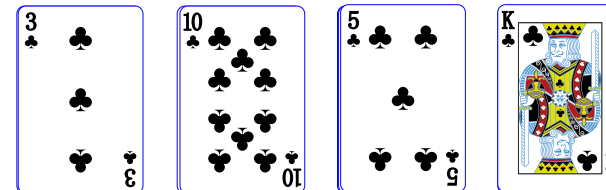
sortie : liste **L** triée dans l'ordre croissant

Pour **i** allant de 2 à **n** :

Si **L(i) < L(i-1)**, alors

L $\leftarrow$ permuter(**L**, **i**, **i-1**)

Sortir : **L**



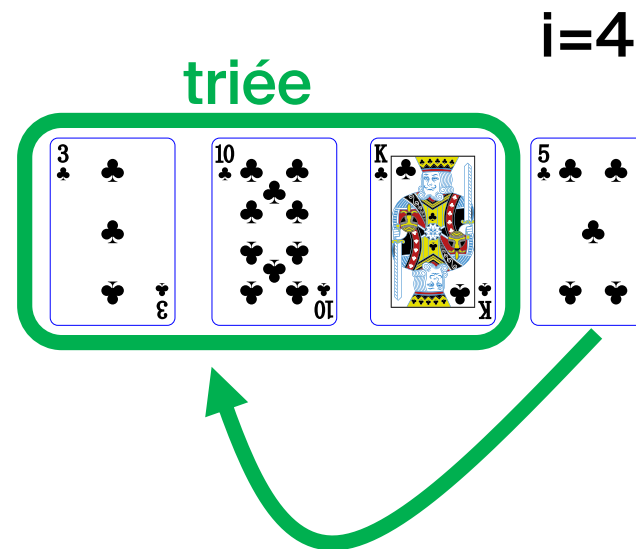
Tri par insertion : algorithme principal



Tri par insertion

entrée : liste L de taille n
sortie : liste L triée dans l'ordre croissant

Pour i allant de 2 à n :
 Si $L(i) < L(i-1)$, alors
 $L \leftarrow \text{insérer}(L, i)$
Sortir : L



Tri par insertion : sous-algorithme 1



insérer

entrée : liste L , indice i

sortie : liste L avec l'élément $L(i)$ bien placé

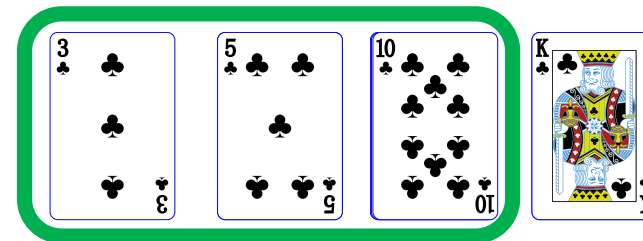
Tant que $i > 1$ et $L(i) < L(i-1)$:

$L \leftarrow \text{permuter}(L, i, i-1)$

$i \leftarrow i-1$

Sortir : L

triée



- **Remarque importante :**
 - Les éléments $L(1) \dots L(i-1)$ doivent être déjà triés pour que ce sous-algorithme fonctionne correctement. Heureusement, c'est le cas ici !

Tri par insertion : sous-algorithme 2



permuter – 1^{er} essai

entrée : liste L , indices j et k
sortie : liste L avec les éléments $L(j)$ et $L(k)$ permutés

$L(j) \leftarrow L(k)$
 $L(k) \leftarrow L(j)$
Sortir : L

permuter

entrée : liste L , indices j et k
sortie : liste L avec les éléments $L(j)$ et $L(k)$ permutés

$temp \leftarrow L(j)$
 $L(j) \leftarrow L(k)$
 $L(k) \leftarrow temp$
Sortir : L

Tri par insertion : algorithme entier



Tri par insertion

entrée : liste L de taille n
sortie : liste L triée dans l'ordre croissant

Pour i allant de 2 à n :
 Si $L(i) < L(i-1)$, alors
 $L \leftarrow \text{insérer}(L, i)$
Sortir : L

insérer

entrée : liste L , indice i
sortie : liste L avec l'élément $L(i)$ bien placé

Tant que $i > 1$ et $L(i) < L(i-1)$:
 $L \leftarrow \text{permuter}(L, i, i-1)$
 $i \leftarrow i-1$
Sortir : L

permuter

entrée : liste L , indices j et k
sortie : liste L avec les éléments $L(j)$ et $L(k)$ permutés

$\text{temp} \leftarrow L(j)$
 $L(j) \leftarrow L(k)$
 $L(k) \leftarrow \text{temp}$
Sortir : L

Question : Tri par insertion

Complétez l'exécution pas à pas de l'algorithme ci-dessous :

Tri par insertion

entrée : liste L de taille n
sortie : liste L triée dans l'ordre croissant

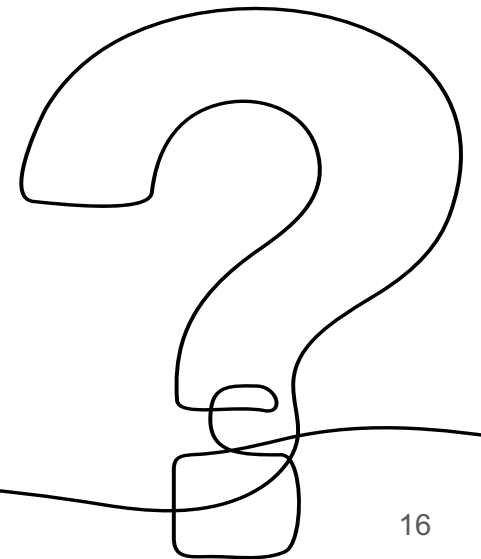
Pour i allant de 2 à n :
 Si $L(i) < L(i-1)$, alors
 $L \leftarrow \text{insérer}(L, i)$
Sortir : L

Exécution pas à pas :

Liste initiale : $L = (5, 2, 8, 1, 9)$

$i = 2$: $L = ($)
 $i = 3$: $L = ($)
 $i = 4$: $L = ($)
 $i = 5$: $L = ($)

Combien de fois le sous-algorithme « insérer » est-il appelé ?



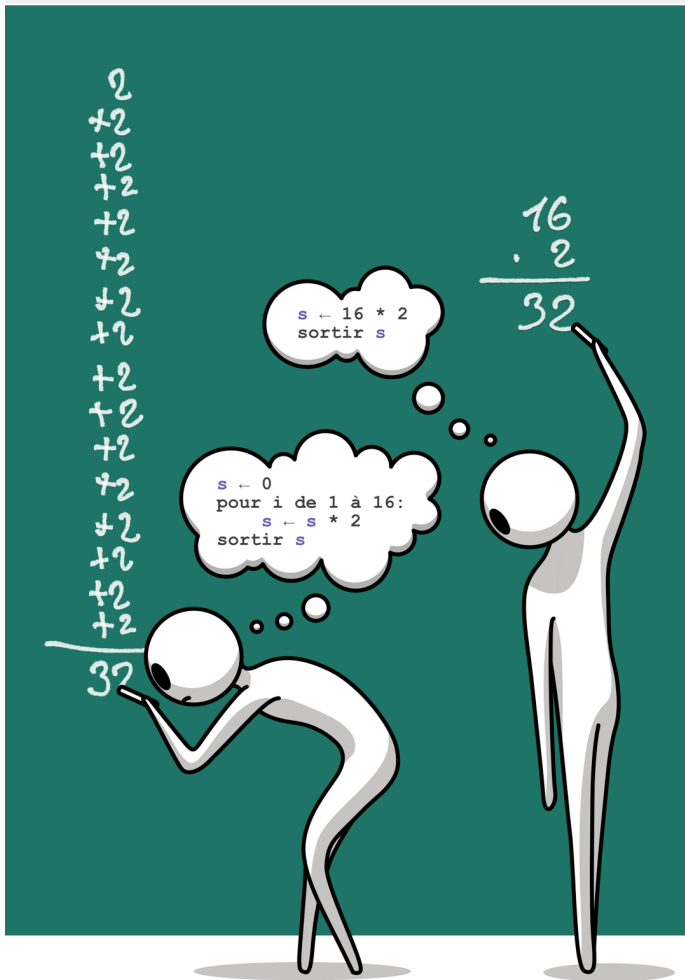
Aujourd'hui

- Sous-algorithmes



- Complexité temporelle
- Notation Grand Theta

Algorithmes : Complexité temporelle



- La complexité temporelle d'un algorithme est son temps d'exécution.
- **Définition plus précise :**
 - ❖ La complexité temporelle d'un algorithme est le nombre **d'opérations élémentaires** effectuées au cours de son exécution, dans le **pire des cas**.
- **opération élémentaire** : addition, soustraction, multiplication ou comparaison
- **pire des cas** : le temps d'exécution peut en effet dépendre des données d'entrée.

Complexité temporelle : Exemples

Algorithme 1

Tant que $1 > 0$:
Afficher "bonjour"

Algorithme 2

entrée : liste L de taille n
sortie : moyenne des n nombres
de la liste

$m \leftarrow 0$
Pour i allant de 1 à n :
 $m \leftarrow m + L(i)$
Sortir : m/n

ICC-T 09 : Tous différents ?

- **Problème à résoudre :**

- Parmi une liste de n objets, identifier si ceux-ci sont tous différents les uns des autres.

Tous différents	
<i>entrée</i> :	liste L de n objets
<i>sortie</i> :	valeur binaire oui/non
$s \leftarrow \text{oui}$ Pour i allant de 1 à $n-1$: Pour k allant de $i+1$ à n : Si $L(i) = L(k)$, alors : Sortir non Sortir : s	

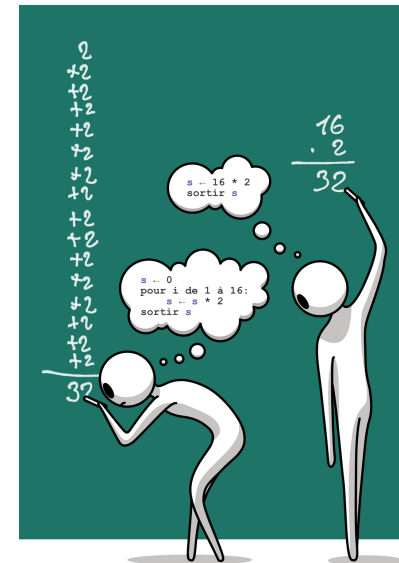
i \ k	1	2	3
1	1,1	1,2	1,3
2	2,1	2,2	2,3
3	3,1	3,2	3,3

$i = 1:$	$k = 2, 3, 4, \dots, n$	$n-1$ comp.
$i = 2:$	$k = 3, 4, \dots, n$	$n-2$ comp.
$i = 3:$	$k = 4, \dots, n$	$n-3$ comp.
$\vdots$	$\vdots$	$\vdots$
$i = n-2:$	$k = n-1, n$	2 comp.
$i = n-1:$	$k = n$	1 comp.

$$1 + 2 + 3 + \dots + (n-2) + (n-1) = \frac{n(n-1)}{2} \text{ comp.}$$

Aujourd'hui

- Sous-algorithmes
- Complexité temporelle
- **Notation Grand Theta**



Complexité temporelle

Valeur minimale

entrée : liste L de nombres entiers, de taille n
sortie : la (ou une des) valeur(s) minimale(s) de la liste

```
 $min \leftarrow L(1)$   
Pour  $i$  allant de 2 à  $n$   
    Si  $L(i) < min$   
         $min \leftarrow L(i)$   
Sortir :  $min$ 
```

Tous différents

entrée : liste L de n objets
sortie : valeur binaire oui/non

```
 $s \leftarrow \text{oui}$   
Pour  $i$  allant de 1 à  $n-1$  :  
    Pour  $k$  allant de  $i+1$  à  $n$  :  
        Si  $L(i) = L(k)$ , alors  $s \leftarrow \text{non}$   
Sortir :  $s$ 
```

Complexité temporelle : comparaison

Valeur minimale $5n - 3$ vs. Tous différents $(5n^2 + n - 2)/2$
En supposant 1 opération = 1 microseconde ($1\mu s$)

n	Valeur minimale $(5n - 3) \times 1\mu s$	Tous différents $(5n^2 + n - 2)/2 \times 1\mu s$	Rapport
1000	$\approx 5\ ms$	$\approx 2,5\ s$	$\times 500$
10 000	$\approx 50\ ms$	$\approx 4,2\ min.$	$\times 5\ 000$
100 000	$\approx 0,5\ s$	$\approx 6,9\ h$	$\times 50\ 000$
1 000 000	$\approx 5\ s$	$\approx 28,9\ jours$	$\times 500\ 000$

Valeurs dominantes n vs. n^2

n	Valeur minimale $n \times 1\mu s$	Tous différents $n^2 \times 1\mu s$	Rapport
1000	$\approx 1\ ms$	$\approx 1\ s$	$\times 1\ 000$
10 000	$\approx 10\ ms$	$\approx 1,7\ min.$	$\times 10\ 000$
100 000	$\approx 100\ ms$	$\approx 2,8\ h$	$\times 100\ 000$
1 000 000	$\approx 1\ s$	$\approx 11,6\ jours$	$\times 1\ 000\ 000$

Notation $\Theta(\cdot)$: introduction

- En général, on évalue la complexité temporelle d'un algorithme en fonction d'un paramètre lié à la taille des données d'entrée (le paramètre n dans les exemples précédents).
- Pourquoi tant s'intéresser à cette complexité temporelle ? Voici un exemple concret :
 - ❖ Supposons qu'un algorithme prenne une minute pour s'exécuter avec des données d'entrée de taille $n = 1'000$. On aimerait savoir en combien de temps (au pire) s'exécutera ce même algorithme avec des données d'entrée de taille $n = 10'000$.
- Si on peut caractériser le nombre d'opérations effectuées par l'algorithme en fonction de n (comme par exemple pour l'algorithme « **Tous différents** » qui effectue $\frac{n(n-1)}{2}$ opérations lors de son exécution, dans le pire des cas, alors on peut répondre à la question ci-dessus.

Notation $\Theta(\cdot)$: définition

- Dans de nombreuses applications, on a affaire à des données d'entrée de grande taille.
- Dans ce cas, on aimerait obtenir des **ordres de grandeur** plutôt que de devoir faire des calculs détaillés.

Soient $f, g : \mathbb{N} \rightarrow \mathbb{R}_+$ deux fonctions non-négatives

On dit que " $f(n)$ est un grand theta de $g(n)$ " et on écrit " $f(n) = \Theta(g(n))$ " s'il existe $0 < C1 < C2 < \infty$ et $N \geq 1$ tels que

$$C1 g(n) \leq f(n) \leq C2 g(n) \quad \text{pour tout } n \geq N$$

Deux exemples :

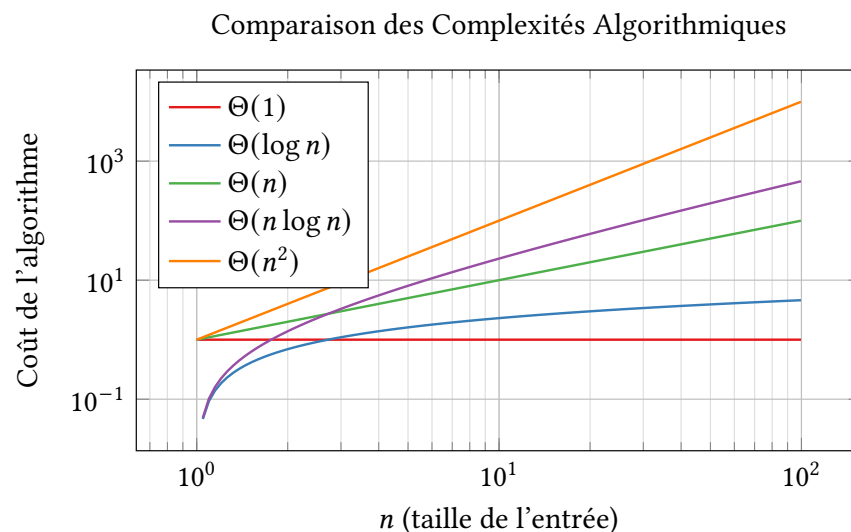
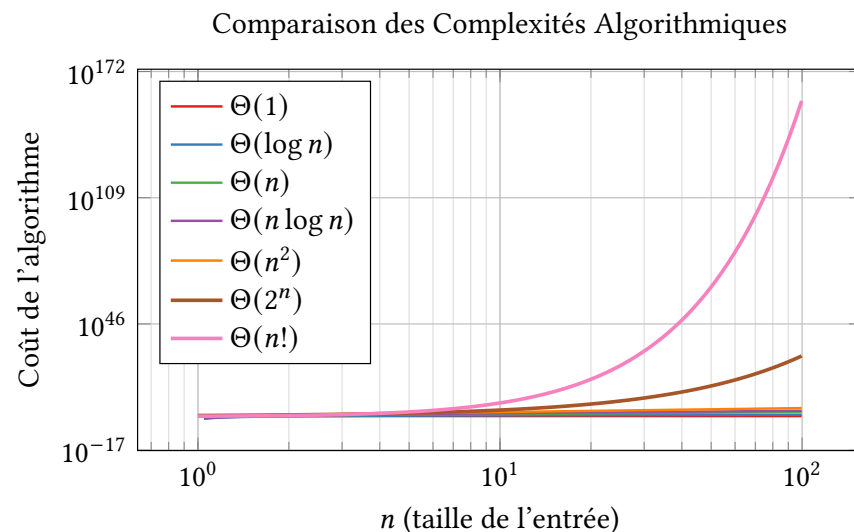
Les fonctions $f(n) = n + 2$ et $f(n) = 3n + 3$ sont **toutes deux** des $\Theta(n)$

La fonction $f(n) = \frac{n(n-1)}{2} + 1$ est un $\Theta(n^2)$

Notation $\Theta(\cdot)$: application

- Revenons à notre exemple :
 - ❖ Supposons qu'un algorithme prenne une minute pour s'exécuter avec des données d'entrée de taille $n = 1'000$. On aimerait savoir en combien de temps (au pire) s'exécutera ce même algorithme avec des données d'entrée de taille $n = 10'000$.
- Si la complexité temporelle de cet algorithme est un $\Theta(n)$ et il prend 1 minute avec une entrée de taille $n=1'000$, alors son temps d'exécution avec $n=10'000$ en entrée vaudra (approximativement) **10 minutes**.
- Si sa complexité temporelle est un $\Theta(n^2)$, alors alors son temps d'exécution avec $n=10'000$ en entrée vaudra (approximativement) **$10 \times 10 = 100$ minutes = 1 h 40 min**.

Notation $\Theta(\cdot)$: Ordres de grandeur



Impraticables : $\Theta(2^n)$, $\Theta(n!)$

Plus lents, mais souvent acceptés : $\Theta(n^2) \dots \Theta(n^k)$, $\Theta(n \cdot \log(n))$

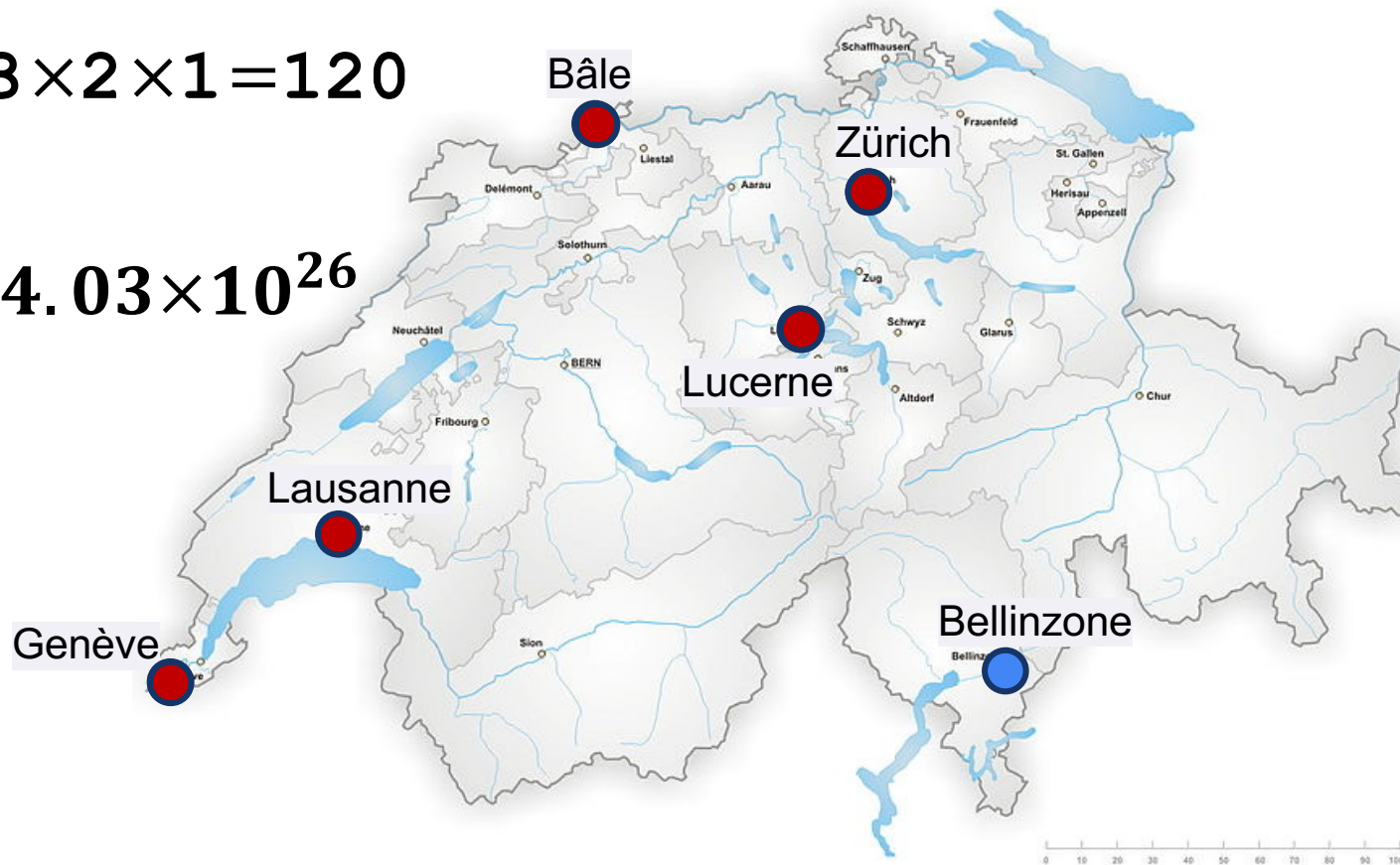
Rapides : $\Theta(1)$, $\Theta(\log n)$, $\Theta(n)$

ICC-T 01 : Problème du voyageur de commerce

$$5 \times 4 \times 3 \times 2 \times 1 = 120$$

$n!$

$$26! \approx 4.03 \times 10^{26}$$



Factorielle	Résultat
1!	1
2!	2
3!	6
4!	24
5!	120
6!	720
7!	5040
8!	40320
9!	362880
10!	3628800
11!	39916800
12!	479001600
13!	6227020800
14!	87178297200
15!	1307674368000
16!	20922789888000
17!	355687428096000
18!	6402373705728000
19!	121645100408832000
20!	2432902008176640000
21!	51090942171709440000
22!	1124000727777607680000
23!	25852016738884976640000
24!	620448401733239439360000
25!	15511210043330985984000000

Notation $\Theta(\cdot)$: Illustration

- Calcul du nombre de paires d'éléments dans l'ensemble $\{ 1, 2, \dots, n \}$
- Pour calculer ce nombre, il existe plusieurs façons de faire :
 - Utilisation de deux boucles imbriquées
 - ❖ Complexité $\Theta(n^2)$
 - Utilisation d'une seule boucle
 - ❖ Complexité $\Theta(n)$
 - Utilisation de la formule mathématique
 - ❖ Complexité $\Theta(1)$

i \ k	1	2	3
1	1,1	1,2	1,3
2	2,1	2,2	2,3
3	3,1	3,2	3,3

```

s ← 0
Pour i allant de 1 à n - 1 :
    Pour j allant de i + 1 à n :
        s ← s + 1
Sortir : s
    
```

```

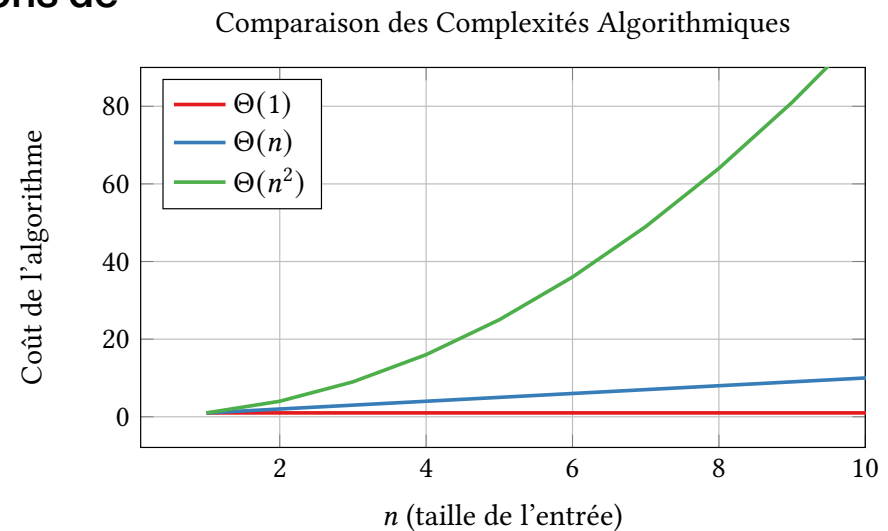
s ← 0
Pour i allant de 1 à n - 1 :
    s ← s + n - i
Sortir : s
    
```

```

s ←  $\frac{n(n-1)}{2}$ 
Sortir : s
    
```

Notation $\Theta(\cdot)$: Illustration

- Calcul du nombre de paires d'éléments dans l'ensemble $\{ 1, 2, \dots, n \}$
- Pour calculer ce nombre, il existe plusieurs façons de faire :
 - Utilisation de deux boucles imbriquées
 - ❖ Complexité $\Theta(n^2)$
 - Utilisation d'une seule boucle
 - ❖ Complexité $\Theta(n)$
 - Utilisation de la formule mathématique
 - ❖ Complexité $\Theta(1)$



Question :

Analysons trois algorithmes différents :

Algorithme A

```
 $s \leftarrow 0$   
Pour  $i$  allant de 1 à  $n$  :  
  Pour  $j$  allant de 1 à  $n$  :  
     $s \leftarrow s + 1$   
Sortir :  $s$ 
```

Algorithme B

```
 $s \leftarrow 0$   
Pour  $i$  allant de 1 à  $n$  :  
  Pour  $j$  allant de  $i$  à  $n$  :  
     $s \leftarrow s + 1$   
Sortir :  $s$ 
```

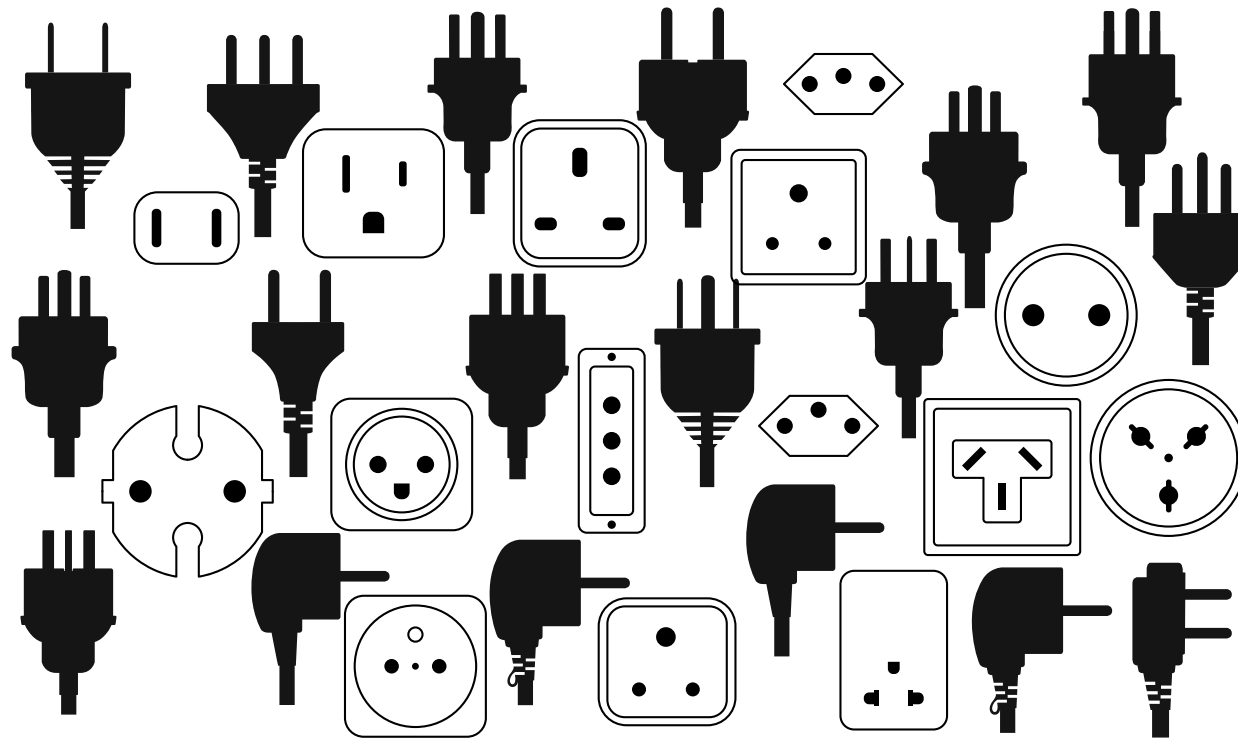
Algorithme C

```
 $s \leftarrow 0$   
 $i \leftarrow 1$   
Tant que  $i \leq n$  :  
   $s \leftarrow s + 1$   
   $i \leftarrow i \times 2$   
Sortir :  $s$ 
```

Quelle est la complexité temporelle Θ de chaque algorithme ?

Si $n = 1000$, combien d'opérations (environ) effectue chaque algorithme ?

Deux font la paire



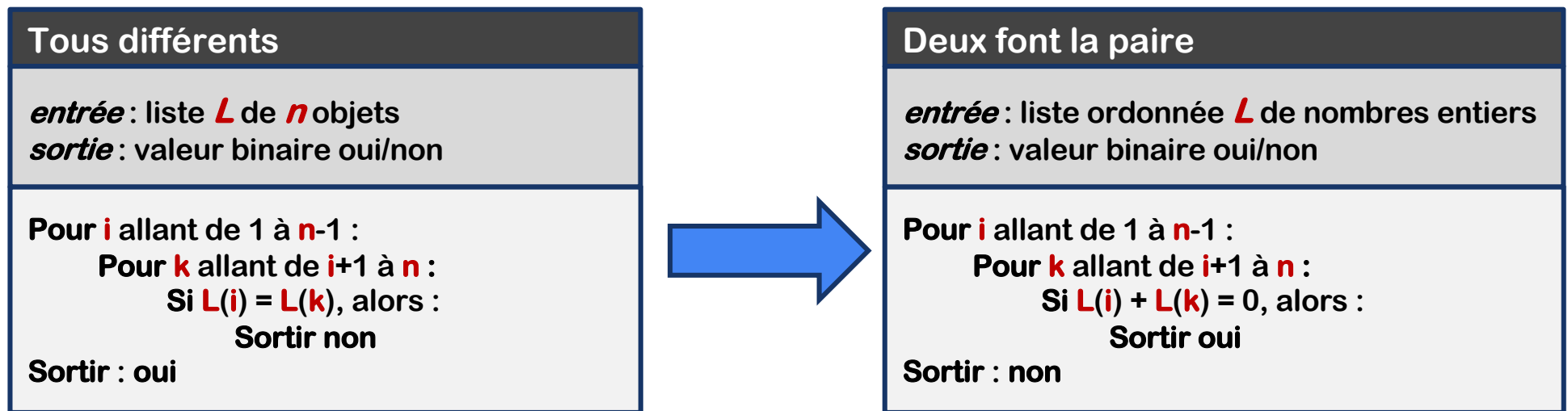
Question : Parmi toutes les fiches et prises ci-dessus, y a-t-il une paire qui s'adapte l'une à l'autre?

Réécriture du problème avec des nombres entiers

- En remplaçant les fiches et les prises par des nombres entiers positifs et négatifs, respectivement, la question précédente se transforme en :
- Etant donnée une liste L de n nombres entiers positifs et négatifs, existe-t-il $i, j \in \{1, \dots, n\}$ tels que $i < j$ et $L(i) + L(j) = 0$?
- **Exemple :**
 - Si $L = (-15, -12, -3, -1, +5, +17, +23)$
alors la réponse est **non**.
 - Si $L = (-14, -3, -1, +3, +7, +10)$,
alors la réponse est **oui**.
- **Note :** Vu que nous avons affaire ici à des nombres entiers, **nous allons supposer de plus que la liste L en entrée est ordonnée.**

Première méthode de résolution

- Etant donnée une liste L de n nombres entiers positifs et négatifs, existe-t-il $i, j \in \{1, \dots, n\}$ tels que $i < j$ et $L(i) + L(j) = 0$?



Première méthode de résolution

Deux font la paire – 1^{er} essai

entrée : liste ordonnée **L** de nombres entiers
sortie : valeur binaire oui/non

```
Pour i allant de 1 à n-1 :  
  Pour k allant de i+1 à n :  
    Si L(i) + L(k) = 0, alors :  
      Sortir oui  
Sortir : non
```

- Les deux boucles imbriquées explorent toutes les paires possibles d'indices $i < j$ dans $\{1 \dots n\}$, qui sont au nombre de

$$(n-1) + (n-2) + \dots + 2 + 1 = \frac{n(n-1)}{2}$$

- donc la complexité temporelle de l'algorithme est $\Theta(n^2)$.
- Question : **Peut-on faire mieux ?**

- Remarque :**
 - L'algorithme précédent n'exploite pas **l'ordre** de la liste L .

Deuxième méthode de résolution

- Etant donnée une liste L de n nombres entiers positifs et négatifs, existe-t-il $i, j \in \{1, \dots, n\}$ tels que $i < j$ et $L(i) + L(j) = 0$?

Deux font la paire

entrée : liste ordonnée L de n nombres entiers
sortie : valeur binaire *oui* / *non*

$i \leftarrow 1$

$j \leftarrow n$

Tant que $i < j$:

Si $L(i) + L(j) = 0$, alors : **Sortir** : *oui*

Si $L(i) + L(j) < 0$, alors : $i \leftarrow i + 1$

Si $L(i) + L(j) > 0$, alors : $j \leftarrow j - 1$

Sortir : *non*

$$L = (-14, -3, -1, +3, +7, +10)$$

- **Remarque** :
 - Complexité temporelle de ce dernier algorithme: $\Theta(n)$ (une seule boucle !).

Question :

Quelle est la complexité temporelle Θ de chaque algorithme ?

Si $n = 10\,000$, combien de fois (environ) la comparaison de somme est-elle effectuée dans le pire des cas ?

Somme cible – Version 1

entrée : liste L de n nombres entiers, cible k
sortie : oui si deux éléments somment à k

```
Pour  $i$  allant de 1 à  $n-1$  :  
  Pour  $j$  allant de  $i+1$  à  $n$  :  
    Si  $L(i) + L(j) = k$ , alors :  
      Sortir oui  
Sortir : non
```

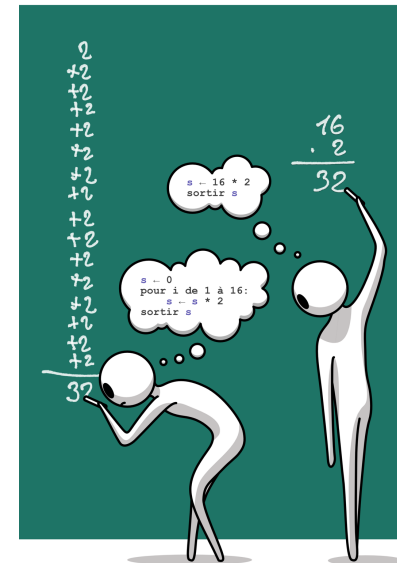
Somme cible – Version 2

entrée : liste ordonnée L de n nombres entiers, cible k
sortie : oui si deux éléments somment à k

```
 $i \leftarrow 1$   
 $j \leftarrow n$   
Tant que  $i < j$  :  
  Si  $L(i) + L(j) = k$ , alors : Sortir : oui  
  Si  $L(i) + L(j) < k$ , alors :  $i \leftarrow i + 1$   
  Si  $L(i) + L(j) > k$ , alors :  $j \leftarrow j - 1$   
Sortir : non
```

Aujourd'hui

- Sous-algorithmes
- Complexité temporelle
- Notation Grand Theta



Θ

Résumé Cours 2 – ICC-T

- Les **sous-algorithmes** permettent de **décomposer** un problème en sous-problèmes plus simples, favorisant ainsi **l'abstraction**, la **réutilisation** du code, une meilleure **lisibilité** et une **maintenance** facilitée des algorithmes.
 - **Tri par insertion**
- La notation **Grand Theta** permet de caractériser précisément **l'ordre de complexité** d'un algorithme.
- Pour un problème donné, il existe souvent **plusieurs algorithmes** de résolution différents.
- En général, des **données d'entrée structurées** permettent une résolution plus efficace du problème.

rafael.pires@epfl.ch

EPFL



Merci