



#### DATA + ALGORITHMS = PROGRAMS

Nicklaus Wirth



#### **Concours**

Et entretiens d'embauche ave une sélection algorithmique « Coding Interview :

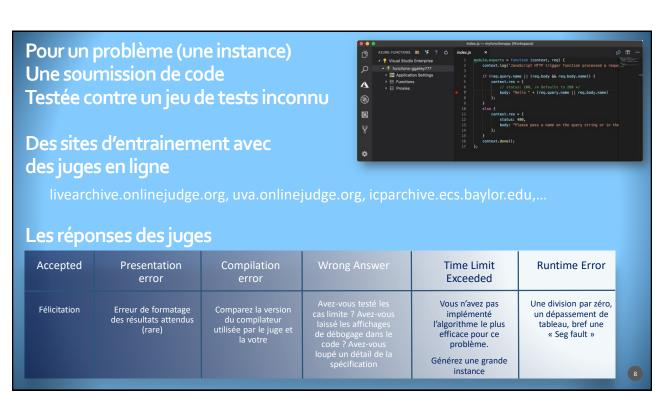
- Depuis une dizaine d'années, des concours de programmation sont organisés dans le monde entier à tous les niveaux.
- Les problèmes sont des variantes de problèmes algorithmiques classiques

4









#### Recrutement

Coding Interview
« Not only code... »

#### 9

#### « Google doesn't look for deep, deep experts » (1/4)

"We would rather hire smart, curious people than people who are deep, deep experts in one area or another," he says, noting that people with strong learning ability can generally find the right answers to unfamiliar questions. "But somebody who's been doing the same thing forever will typically just replicate what they've seen before." - <a href="https://www.businessinsider.com/qualities-google-looks-for-in-job-candidates-2014-4?r=US&IR=T">https://www.businessinsider.com/qualities-google-looks-for-in-job-candidates-2014-4?r=US&IR=T</a>

#### Google does want people with high "cognitive ability."

"If you hire someone who is bright, and curious, and can learn, they're more likely to come up with a new solution that the world hasn't seen before," bock explained in a google+ q&a. "This looking for cognitive ability stems from wanting people who are going to reinvent the way their jobs are going to work rather than somebody who's going to come in and do what everybody else does. We recruit for aptitude, for the ability to learn new things and incorporate them."

#### Google seeks out people with "grit."

Bock spoke with the times about a time he was on a campus talking to a student double-majoring in computer science and math. The student was thinking about switching out of computer science — it was too difficult.

"I told that student they are much better off being a b student in computer science than an a+ student in english," he recalls. Taking computer science "signals a rigor in your thinking and a more challenging course load. That student will be one of our interns this summer."

the ability to keep slogging through difficult work — is more important for success than raw iq.

#### « Google doesn't look for deep, deep experts » (1/2 oops... 2/4)

#### Google wants to know whether candidates can tackle difficult projects.

Google's interviews include questions about the candidate's concrete experiences, starting with queries like "give me an example of a time when you solved an analytically difficult problem."

By asking people to speak of their own experiences, bock says, you get two kinds of information: "you get to see how they actually interacted in a real-world situation, and the valuable 'meta' information you get about the candidate is a sense of what they consider to be difficult."

#### Google wants candidates with analytical skills.

Basic computer science skills will do, bock says, since they signal "the ability to understand and apply information" and think in a formal, logical, and structured way. But there are options beyond CS. Bock says that taking statistics while he was in business school was "transformative" for his career.

"Analytical training gives you a skill set that differentiates you from most people" he says.

#### Google expects people to meet ridiculously high standards.

"We don't compromise our hiring bar, ever," bock says. Because of this, job listings stay open longer at google than you'd expect, he says — they have to kiss a lot of frogs before finding the one. Google doesn't care about Top gpas (Grade Point AverageS). They'll be zealots about their point of view.

#### « Google doesn't look for deep, deep experts » (3/4)

#### Gpas (Grade Point AverageS) and test scores don't correlate with success at the company.

"Academic environments are artificial environments. People who succeed there are sort of finely trained; they're conditioned to succeed in that environment," bock says.

While in school, people are trained to give specific answers. "It's much more interesting to solve problems where there isn't an obvious answer," bock says. "You want people who like figuring out stuff where there is no obvious answer."

#### Google wants to know how much candidates have accomplished compared to their peers.

When bock was explaining how to write resumes to Thomas Friedman at the times, he said that most people miss that the formula for writing quality resumes is simple: "i accomplished x, relative to y, by doing z."

For example, bock explained that a lot of people would just write, "i wrote editorials for the new york times."

But a stand-out resume would be more specific about their accomplishments and how they compared to others. Bock gives a better example: "had 50 op-eds published compared to average of 6 by most op-ed [writers] as a result of providing deep insight into the following area for three years."

#### « Google doesn't look for deep, deep experts » (4/4)

#### Google looks for employees who know when to step up and take a leadership role.

Bock doesn't care for "traditional leadership." - "We don't care," he insists. "What we care about is, when faced with a problem and you're a member of a team, do you, at the appropriate time, step in and lead. And just as critically, do you step back and stop leading, do you let someone else? Because what's critical to be an effective leader in this environment is you have to be willing to relinquish power."

#### Google wants to see people who take ownership of projects.

With that sense of ownership, you'll feel responsible for the fate of a project, making you ready to solve any problem. But you also need to defer when other people have better ideas: "your end goal," explained bock, "is what can we do together to problem-solve. I've contributed my piece, and then I step back."

#### Google wants to see humility, too.

You need "intellectual humility" to succeed at google. "Without humility, you are unable to learn." This is a common problem among the well-educated; elite school grads tend to plateau. Successful folks don't often experience failure. So they don't know how to learn from failure. Instead of having an opportunity to learn, they blame others. Bock explains: they, instead, commit the fundamental attribution error, which is if something good happens, it's because i'm a genius. If something bad happens, it's because someone's an idiot or I didn't get the resources or the market moved. ... But people who are the most successful here, who we want to hire, will show humility.

#### Le Hacker Rank (par pays... 1/2) - évaluation des développeurs :

The study looked at the results of 1.4 million of hackerrank's coding test submissions, called "challenges," during the last few years.

"According to our data, china and russia score as the most talented developers.



Chinese programmers outscore all other countries in mathematics, functional programming, and data structures challenges, while Russians dominate in algorithms, the most popular and most competitive arena," said Ritika Trikha, a blogger at hacker rank.

United States and India provide the majority of competitors on hacker rank but only manage to rank 28th and 31st, respectively. "If we held a hacking olympics today, our data suggests that China would win the gold, Russia would take home a silver, and Poland would nab the bronze," Trikha said. "Though they certainly deserve credit for making a showing, the United States and India have some work ahead of them before they make it at least into the top 25."

#### Une évaluation des développeurs de différents pays (2/2)

Hacker rank's coding challenges cover aspects of computing ranging from languages to algorithms, security and distributed systems. Developers are scored based on a combination of accuracy and speed. The algorithms category has nearly 40 percent of developers competing, featuring tests on sorting data, dynamic programming, keyword searches and other logic-based tasks. Following algorithms were java and data structure tests, with 10 percent of developers participating. Distributed systems and security were the least popular tests, although thousands still took them.

To determine which nation had the highest-scoring programmers, HackerRank looked at each country's average score across domains. Data was restricted to the top 50 countries with the most developers on HackerRank. Following China (1) and Russia(2) with the top developers were Poland (3), Switzerland (4), Hungary (5), Japan (6), Taiwan (7), France (8), Czech republic (9), and Italy (10).

The 100 score represents the country's being first in the rankings (China). "But China only won by a hair. Russia scored 99.9 out of 100, while Poland and Switzerland round out the top rankings with scores near 98. Pakistan scores only 57.4 out of 100 on the index, (ranking 50th)."

Poland was tops in java testing, France led in C++, Hong Kong in Python, Japan in artificial intelligence, and Switzerland in databases. Ukrainian programmers led in security, while Finland was top in ruby coding challenges.



#### La complexité

Pour écrire un programme efficace, il faut un algorithme de bonne complexité face au problème.

La complexité d'exprime en fonction de la taille de l'entrée ou un paramètre de l'entrée 'n'

Pour comparer les complexités on utilise la notation de Landau.

**Hiérachie**, du terme le plus important au moins important (liste non-exhaustive) (suite)

- Exponentiel en n : O(2<sup>n</sup>)
- Polynomial en n : O(n<sup>k</sup>)
- Un terme logarithmique, c'est-à-dire de la forme c x log<sub>2</sub>(n) où c ne dépend pas de n. On le notera O(log(n)). Exemple: 6 log<sub>2</sub>(n) opérations.
- Un terme constant, qui ne dépend pas de n. On le notera O(1). Exemple: 45 opérations.

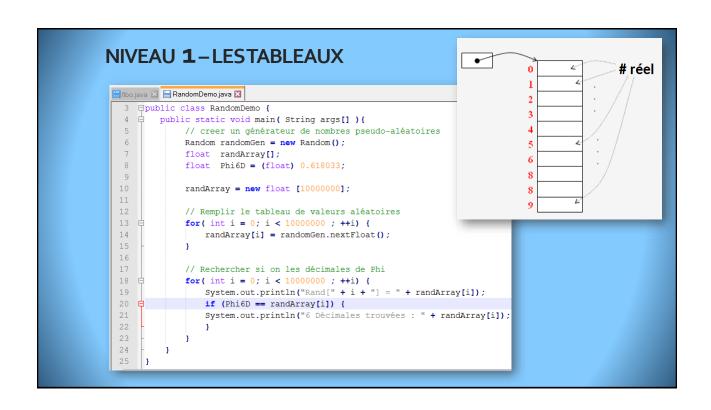


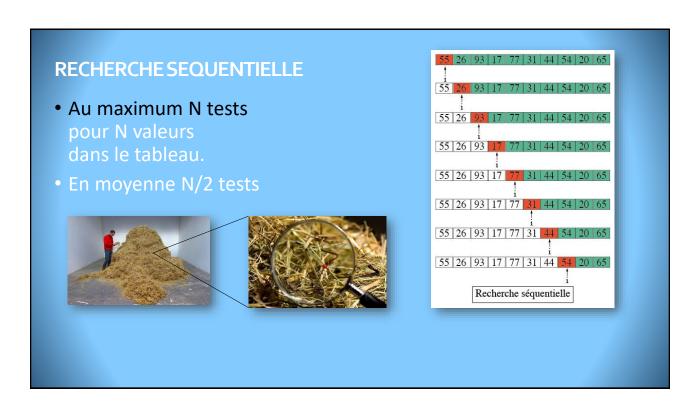


Ordre de grandeur du temps nécessaire à l'exécution d'un algorithme d'un type de complexité										
Temps	Type de complexité	Temps pour n = 5	Temps pour n = 10	Temps pour n = 20	Temps pour n = 50	Temps pour n = 250	Temps pour n = 1 000	Temps pour n = 10 000	Temps pour n = 1 000 000	Problème exemple
O(1)	complexité constante	10 ns	10 ns	10 ns	10 ns	10 ns	10 ns	10 ns	10 ns	accès à une cellule de tableau
$O(\log(n))$	complexité logarithmique	10 ns	10 ns	10 ns	20 ns	30 ns	30 ns	40 ns	60 ns	recherche dichotomique
$O(\sqrt{n})$	complexité racinaire	22 ns	32 ns	45 ns	71 ns	158 ns	316 ns	1 µs	10 µs	test de primalité naïf
O(n)	complexité linéaire	50 ns	100 ns	200 ns	500 ns	2.5 µs	10 µs	100 µs	10 ms	parcours de liste
$O(n\log(n))$	complexité linéarithmique	40 ns	100 ns	260 ns	850 ns	6 µs	30 µs	400 μs	60 ms	tris par comparaisons optimaux (comme le tri fusion ou le tri par tas)
$O(n^2)$	complexité quadratique (polynomiale)	250 ns	1 µs	4 µs	25 µs	625 µs	10 ms	1 s	2.8 heures	parcours de tableaux 2D

$O(n^2)$	complexité quadratique (polynomiale)	250 ns	1 µs	4 µs	25 µs	625 µs	10 ms	1 s	2.8 heures	parcours de tableaux 2D
$O(n^3)$	complexité cubique (polynomiale)	1.25 µs	10 µs	80 µs	1.25 ms	156 ms	10 s	2.7 heures	316 ans	multiplication matricielle naïve
$2^{\mathrm{poly}(\log(n))}$	complexité sous- exponentielle	30 ns	100 ns	492 ns	7 μs	5 ms	10 s	3.2 ans	$10^{20}$ ans	factorisation d'entiers avec GNFS (le meilleur algorithme connu en 2018)
$2^{\mathrm{poly(n)}}$	complexité exponentielle	320 ns	10 µs	10 ms	130 jours	10 <sup>59</sup> ans				problème du sac à dos par force brute
O(n!)	complexité factorielle	1.2 µs	36 ms	770 ans	10 <sup>48</sup> ans					problème du voyageur de commerce avec une approche naïve
$2^{2^{\mathrm{poly(n)}}}$	complexité doublement exponentielle	4.3 s	10 <sup>278</sup> ans							décision de l'arithmétique de Presburger

I	Computer Science								
ı	2 <sup>n</sup> (Level o) r		n² (Level 1)	n (Level 2)	log(n) (Level 3)	Comments			
	data structures	Doesn't know the difference between Array and LinkedList	Able to explain and use Arrays, LinkedLists, Dictionaries etc in practical programming tasks	Knows space and time tradeoffs of the basic data structures, Arrays vs LinkedLists, Able to explain how hashtables can be implemented and can handle collisions, Priority queues and ways to implement them etc.	Knowledge of advanced data structures like B-trees, binomial and fibonacci heaps, AVL/Red Black trees, Splay Trees, Skip Lists, tries etc.				
	algorithms  Unable to find the average of numbers in an array (It's hard to believe but I've interviewed such candidates)		Basic sorting, searching and data structure traversal and retrieval algorithms	Tree, Graph, simple greedy and divide and conquer algorithms, is able to understand the relevance of the levels of this matrix.	Able to recognize and code dynamic programming solutions, good knowledge of graph algorithms, good knowledge of numerical computation algorithms, able to identify NP problems etc.	Working with someone who has a good topcoder ranking would be an unbelievable piece of luck!			
	systems programming	Doesn't know what a compiler, linker or interpreter is	Basic understanding of compilers, linker and interpreters. Understands what assembly code is and how things work at the hardware level. Some knowledge of virtual memory and paging.	Understands kernel mode vs. user mode, multi-threading, synchronization primitives and how they're implemented, able to read assembly code. Understands how networks work, understanding of network protocols and socket level programming.	Understands the entire programming stack, hardware (CPU + Memory + Cache + Interrupts + microcode), binary code, assembly, static and dynamic linking, compilation, interpretation, JIT compilation, garbage collection, heap, stack, memory addressing				







Le maximum N tests est l'ordre de grandeur du temps d'attente maximal Avec beaucoup de chance 1 tests ©!



Rand[249300] = 0.66389877 Rand[249301] = 0.0677259 Rand[249302] = 0.6618339 Rand[249303] = 0.2822563 Rand[249304] = 0.030273736 Rand[249305] = 0.21487772 Rand[249306] = 0.6430948 Rand[249307] = 0.016756117 Rand[249308] = 0.7717602 Rand[249309] = 0.3684681 Rand[249310] = 0.4684795 Rand[249311] = 0.17858976 Rand[249312] = 0.1234051 Rand[249313] = 0.4094249 Rand[249314] = 0.16900414 Rand[249315] = 0.75723165Rand[249316] = 0.18286616 Rand[249317] = 0.13986063 Rand[249318] = 0.18248314 Rand[249319] = 0.5175898 Rand[249320] = 0.16588879 Rand[249321] = 0.49524796 Rand[249322] = Rand[249323] = 0.07022601 Rand[249324] = 0.63770694 Rand[249325] = 0.100140214 and[249326] = 0.75005656 Rand[249327] = 0.959113 Rand[249328] = 0.90208834

#### NIVEAU 2 - RECHERCHE DITE DICHOTOMIQUE

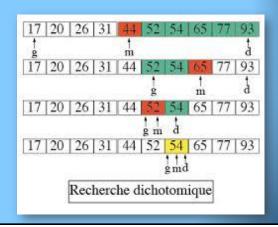
Le maximum de tests sera de l'ordre de log2(N) Ex: pour 4 milliards de valeurs, 32 comparaisons au maximum...

- (1) Le tableau doit être trié (ordre croissant par exemple)
- (2) On regarde au milieu du tableau

<u>Si</u> l'élément du milieu est plus petit, <u>Alors</u> on va regarder dans le sous tableau à droite

Sinon on va regarder dans le sous tableau à gauche...

Finsi



## NIVEAU 3 PREPARATION D'UN QUICKSORT

Un tri sophistiqué appelé – quicksort Complexité moyenne : n.Log(n) Linéarithmique / quasi-linéaire

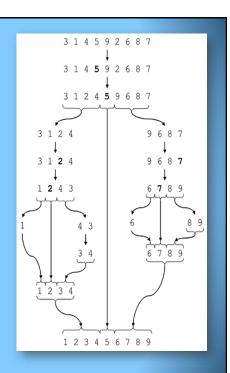
#### **Principe:**

On choisit une valeur « pivot »

Toutes les valeurs au dessus du pivot passent dans le sous tableau à droite.

Toutes les valeurs en dessous passent dans le sous tableau à gauche.

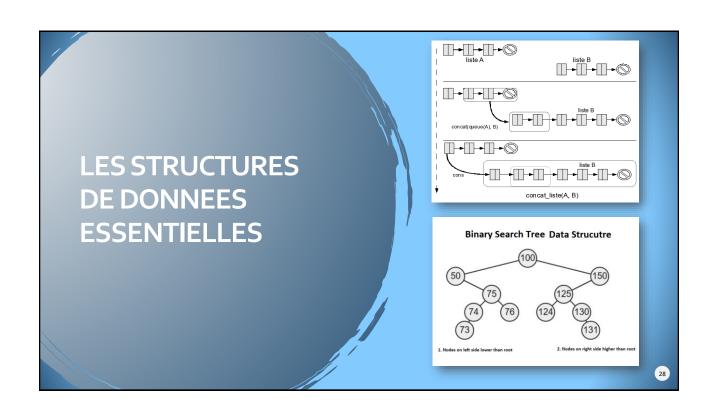
On recommence pour chaque sous tableau

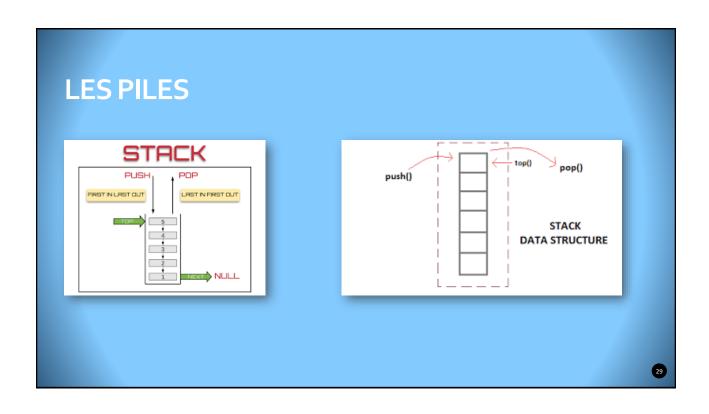


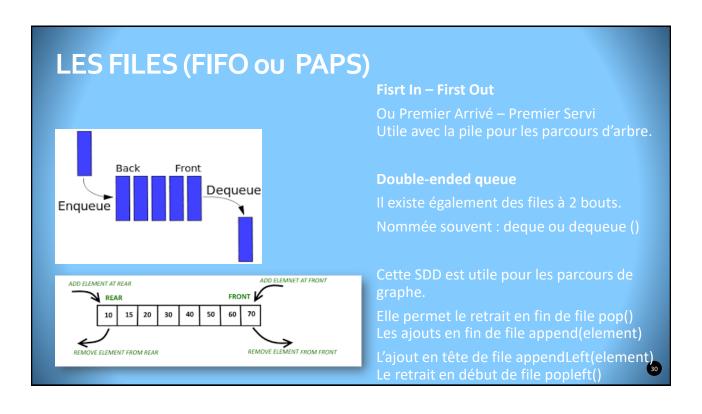
```
NIVEAU 3-EXTRAIT DE CODE JAVA-QUICKSORT
                                                                                 RFCURSIF
public static int Partition(int[] numbers, int left, int right)
   int pivot = numbers[left];
                                                  public static void main(String[] args)
   while (true)
                                                      int[] numbers = { 3, 8, 7, 5, 2, 1, 9, 6, 4 };
       while (numbers[left] < pivot) left++;</pre>
       while (numbers[right] > pivot) right--;
                                                      System.out.println("QuickSort By Recursive Method");
       if (left < right) {</pre>
           int temp = numbers[right];
                                                      QuickSort Recursive (numbers, 0, len - 1);
           numbers[right] = numbers[left];
                                                      for (int i = 0; i < 9; i++) {
           numbers[left] = temp;
                                                          System.out.println(numbers[i]);
                          public static void QuickSort Recursive(int[] arr, int left, int right)
           return right;
                                 if(left < right)</pre>
                                     int pivot = Partition(arr, left, right);
                                     if(pivot > 1) QuickSort Recursive(arr, left, pivot - 1);
                                     if(pivot + 1 < right) QuickSort Recursive(arr, pivot + 1, right);</pre>
```

## Synthèse en français des principales complexités algorithmique <a href="http://fr.cppreference.com/w/cpp/complexity">http://fr.cppreference.com/w/cpp/complexity</a>

Complexité	Vitesse	Description	Formulation	Exemple
Factorielle	très lent	temps d'exécution proportionnel à N <sup>N</sup>	N!	Résolution par recherche exhaustive du problème du voyageur de commerce.
Exponentielle	lent	temps d'exécution proportionnel à une valeur donnée à la puissance N	KN	Résolution par recherche exhaustive du Rubik's Cube.
Polynomiale	moyen	temps d'exécution proportionnel à N à une puissance donnée	NK	Tris par comparaison, comme le tri à bulle (N <sup>2</sup> ).
Quasi-linéaire	assez rapide	temps d'exécution intermédiaire entre linéaire et polynomial	N * log(N)	Tris quasi-linéaires, comme le Quicksort.
Linéaire	rapide	temps d'exécution proportionnel à N	N	Itération sur un tableau.
Logarithmique	très rapide	temps d'exécution moyen proportionnel au logarithme de N	log(N)	Recherche dans un arbre binaire.
Constante	le plus rapide	temps d'exécution donné, quel que soit le nombre d'éléments	1	Recherche par index dans un tableau.



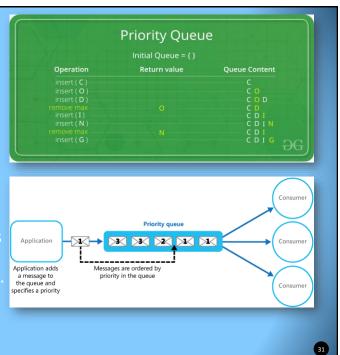


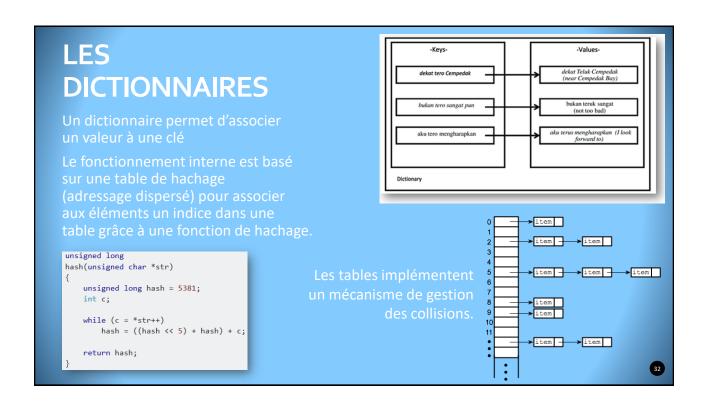


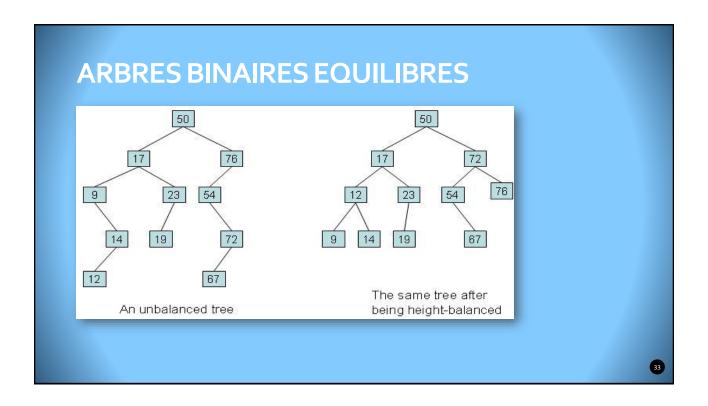


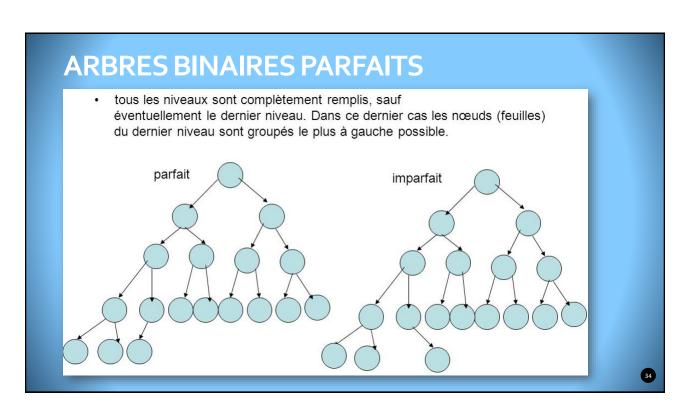
Une file de priorité est un type abstrait permettant d'ajouter des éléments et de retirer un élément de clé ordonné.

Utile dans la construction d'un code de compression style « Huffman » ou dans la recherche d'un plus cours chemin entre deux nœuds d'un graphe (algorithme de Dijkstra).









#### **LESTAS**

Un tas (heap) en anglais est appelé arbre tournoi. C'est un arbre vérifiant la propriété de « tas » - il est partiellement trié.

Les fils de chaque nœud ont une clé plus grande de le nœud racine pour une certaine relation d'ordre (Min Heap)

Il existe une variante ou les fils ont tous une clé plus petite (Max Heap)

Quand ces arbres sont binaires, on peut extraire l'élément le plus petit immédiatement et insérer des éléments avec un coût logarithmique.

35

## 

## LESTECHNIQUES UTILES

- Comparer des éléments
- Trier
- Balayage (parcours de gauche à droite avec un traitement spécifique à chaque élément)
   for\_each
- Algorithmes gloutons établissement de solution pas à pas avec un choix de maximisation de critère local (voir structures combinatoires de type matroïdes)
- Programmation dynamique décomposer un problème en sous-problèmes et trouver la solution optimale du problème principal à partir des solutions aux sousproblèmes.

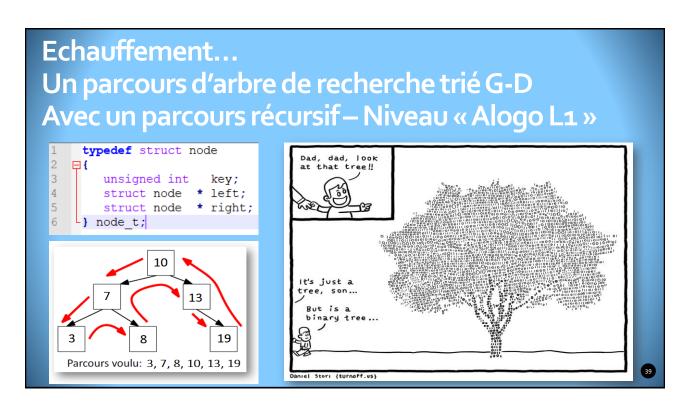
#### **Array Sorting Algorithms** Algorithm Time Complexity Space Complexity Worst Worst Best Average $\Omega(n \log(n))$ Θ(n log(n)) 0(log(n)) Quicksort Mergesort $\Omega(n \log(n))$ Θ(n log(n)) 0(n log(n)) 0(n) Θ(n log(n)) <u>Timsort</u> 0(n log(n)) Heapsort $\Omega(n \log(n))$ Θ(n log(n)) O(n log(n)) Bubble Sort Ω(n) 0(n^2) 0(n^2) **Insertion Sort** 0(n^2) Ω(n) Selection Sort 0(n^2) Tree Sort $\Omega(n \log(n))$ $\theta(n \log(n))$ 0(n^2) Shell Sort $\Omega(n \log(n)) \Theta(n(\log(n))^2)$ O(n(log(n))^2) Bucket Sort Radix Sort Counting Sort Cubesort 0(n log(n))

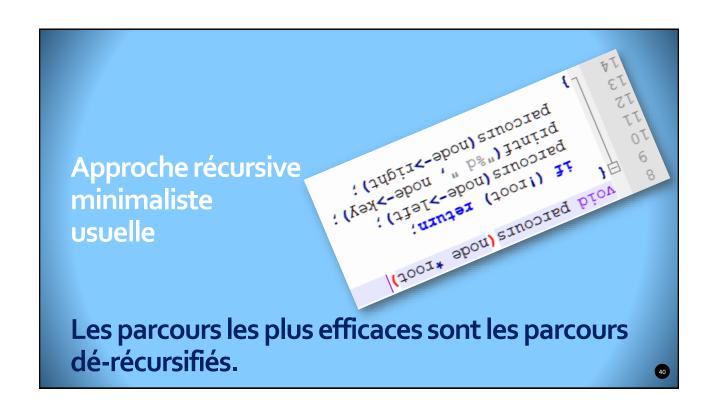
## AUTRES TECHNIQUES UTILES

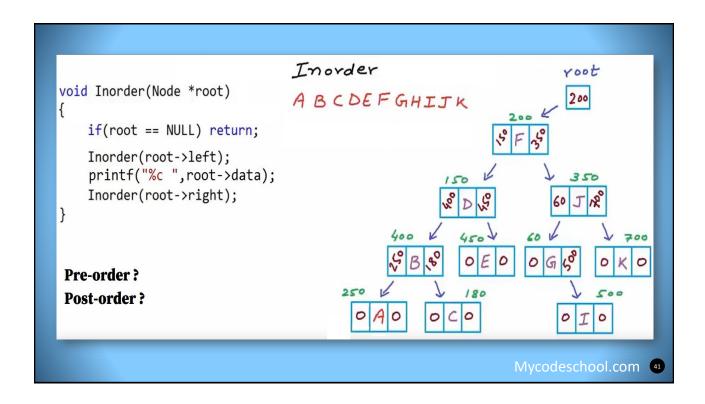
- Coder des ensembles dans des entiers
- Partages équitables en 3 parties
- Recherche dichotomique
- Recherche sans bornes supérieures connues
- Inverser une fonction
- Recherche trichotomiques
- Recherche dans un intervalle [0.. 2<sup>k</sup>]

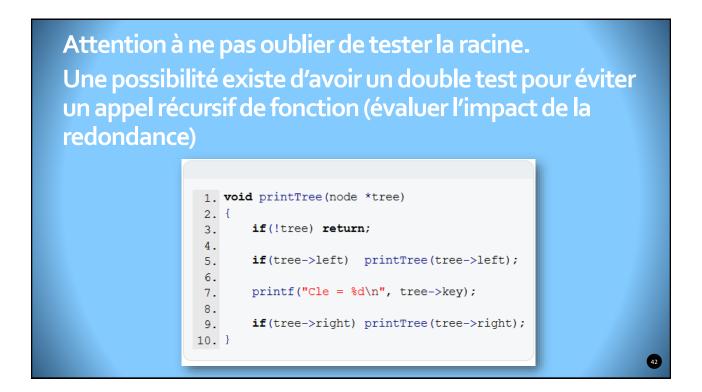
• ..

# Les algorithmes R'U'F'U FR FR U R'U'F' FR U R'U'F'







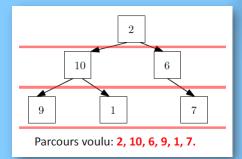


## Rappel L1-S2: parcours d'arbre binaire de recherche en largeur avec un file.

```
Proc parcoursEnLargeur(nœud racine):

file F:=NewFile()
enfiler(F, racine)
nœud N

Tant que non(estVide(F)):
N:=valeurDebut(F)
print(N.data)
defiler(F)
Si N.filsG=! -1: /* si N a un fils gauche */
enfiler(F, N.filsG)
Si N.filsD!=-1:
enfiler(F, N.filsD)
```



Idem pour parcours d'arbre binaire en profondeur (dé-récursifié). Cette fois il faut une pile.

```
Proc parcoursEnProfondeur(nœud racine):

pile P:=NewPile()

empiler(P, racine)

nœud N

Tant que non(estVide(P)):

N:=valeurTop(P)

print(N.data)

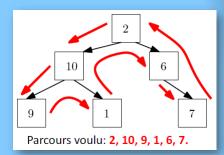
depiler(P)

Si N.filsG=! -1: /* si N a un fils gauche */

empiler(P, N.filsG)

Si N.filsD!=-1:

empiler(P, N.filsD)
```



## MERCI DE VOTRE ATTENTION