15. IFT2015 Structures de données: Liste détaillée de sujets1

Introduction

LE BUT DE CE DOCUMENT est de définir les compétences et connaissances requises dans le cours IFT2015 à l'examen final. L'examen constitue également la deuxième partie de l'examen pré-doctorale en structures de données.

- Les sujets marqués par ★★ correspondent plutôt à un niveau «A+/A».
- ★ Les notes marginales sont des références aux ouvrages suivants S Sedgewick, R. Algorithmes en Java, 3e édition (2004) SW Sedgewick, R. et K. Wayne. Algorithms, 4e édition (2011)
- * Les notes de cours et des liens vers des articles Wikipedia sont affichés sur le site http://ift2015a16.wordpress.com/.
- * Aucune documentation ne sera permise à l'examen final.

¹ Detailed List of Subjects for the Final Examination — English translation starts on Page 10





Principes d'analyse d'algorithmes

Références

- ⊳ Sedgewick chapitre 2
- ▶ Sedgewick & Wayne §1.4²
- Notes sur les fondations : notes01-recursion.pdf.
- Notes sur l'analyse d'algorithmes : notes04-analysis.pdf.

Sujets

- * Principes de base : pire cas, meilleur cas, moyen cas.
- * Croissance de fonctions communes : constantes, logarithmiques, polynomiales, exponentielles. Factorielle(n!), approximation de Stirling³ nombres Fibonacci⁴, nombres harmoniques⁵, logarithme itéré.
- * Notation asymptotique⁶: définitions de grand O(f), petit o(f), $\Theta(f)$ et $\Omega(f)$.

Asymptotiques exactes $f \sim g$. Expressions avec O() ou o(), règles d'arithmétique : O(f) + O(g), $O(f) \cdot O(g)$. Relations avec la limite

$$\lim_{n \to \infty} \frac{f(n)}{g(n)} = c > 0 \qquad \Rightarrow \qquad f(n) = O(g(n));$$

$$\lim_{n \to \infty} \frac{f(n)}{g(n)} = 0 \qquad \Leftrightarrow \qquad f(n) = o(g(n));$$

$$\lim_{n \to \infty} \frac{f(n)}{g(n)} = 1 \qquad \Leftrightarrow \qquad f(n) \sim g(n)$$

- * Application de la définition pour démontrer f = O(g) ou f = o(g).
- * Détermination du temps de calcul et d'usage de mémoire pour algorithmes (itératifs) simples, et pour algorithmes récursifs (comme expression récursive).
- * Récurrences simples.

$$f(n) = f(n-1) + O(1)$$
 $f(n) = O(n);$
 $f(n) = f(n/2) + O(1)$ $f(n) = O(\log n);$
 $f(n) = 2f(n/2) + O(1)$ $f(n) = O(n\log n);$
 $f(n) = 2f(n/2) + O(n)$ $f(n) = O(n\log n);$

- ** Preuve par induction pour récurrences asymptotiques.
- * Notion de temps amorti.
- ** Preuves de résultats sur le coût amorti d'opérations. Principe d'analyse crédit/débit⁷.
- * Validation expérimentale de temps de calcul

² http://algs4.cs.princeton.edu/14analysis/

S§2.1,2.2,2.7 S§2.3 3 $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^{n}$ $^{4}F_{0}=0;F_{1}=1;F_{n}=F_{n-1}+F_{n-2}\{n>$ $^{5}H_{n} = \sum_{i=1}^{n} 1/i = \ln n + \gamma + o(1)$ ⁶ W_(fr):comparaison asymptotique S§2.4

S§2.5,2.6

⁷ W_(en):accounting method

Structures élémentaires et types abstraits

Références

- ⊳ Sedgewick chapitres 3 et 4
- \triangleright Sedgewick & Wayne §1.1⁸, §1.2⁹, §1.3¹⁰
- Notes sur la liste chaînée : notes02-linkedlist.pdf.
- Notes sur les tableaux : notes03-tableaux.pdf.

Sujets

- * Blocs de construction pour programmes Java.
- * Notions de type abstrait, interface, implantation, client.
- * Types abstraits de files généralisées, piles et queues/files FIFO.
- ★ Listes chaînées¹¹. Variations : listes circulaires, doublement chaînées. Sentinelles¹² pour la tête et/ou la queue. Manipulation d'éléments sur la liste, insertion et suppression. Parcours d'une liste.
- \star Tableaux¹³.
- * Implantations de pile et de queue par tableaux ou listes chaînées. Efficacité d'implantations différentes (temps de calcul pour les opérations standardes). Débordement.

- 8 http://algs4.cs.princeton.edu/11model/
- 9 http://algs4.cs.princeton.edu/12oop/
- 10 http://algs4.cs.princeton.edu/13stacks/

S§3.1;SW§1.1

S§4.1;SW§1.2

S§4.2,4.7

11 W_(fr):liste chaînée

S§3.3,3.4

 12 $W_{(en)}$:sentinel

 13 W_(fr):tableau

S§4.4,4.5,4.7;SW§1.3

Tableaux de hachage

Références

- ⊳ Sedgewick chapitre 14
- Sedgewick & Wayne §3.4¹⁴, §3.5¹⁵.
- Notes sur le hachage : notes06-hashing.pdf.

- 14 http://algs4.cs.princeton.edu/34hash/
- 15 http://algs4.cs.princeton.edu/

35applications/

Sujets

* Notions de base pour tableaux de hachage 16 : facteur de charge/remplissage, S§14.1 collisions.

16 W_(fr):table de hachage

- * Fonctions de hachage : méthodes de la division et de la multiplication.
- * Résolution de collisions par chaînage séparé. Coût moyen des opérations de l'interface (table de symboles) en fonction de la facteur de charge.
- * Addressage ouvert : notion de sondage/test. Procédures de recherche et d'insertion avec addressage ouvert. Suppression paresseuse et hachage dynamique. Sondage linéaire, grappe forte. Double hachage.

.....

S§14.2

S§14.3-14.6

Appartenance-union

Références

- ▶ Sedgewick §1.2–1.3
- ▷ Sedgewick & Wayne §1.5¹⁷
- Notes sur Union-Find : notes07-unionfind.pdf.

17 http://algs4.cs.princeton.edu/15uf/

Sujets

* Problème de connexité, opérations d'appartenance-union.

★ Structure Union-Find¹⁸. Astuces: union-par-rang/union-par-taille, compression de chemin.

** Coût amorti d'opérations : $O(\alpha(m, n))$ pour Union-Find avec union équilibrée et compression de chemin; fonction d'Ackermann¹⁹ et son inverse.

S§1.2

18 W(fr):union-find S§1.3;SW§1.5

19 W_(fr):fonction d'Ackermann

Arbres

Références

- ▶ Sedgewick §4.3, §5.4–5.7
- Notes sur les arbres : notes08-trees.pdf.

Sujets

*	Terminologie pour structures arborescentes : arbre k-aire, hauteur,	S§5.4
	niveau, profondeur. Implémentation d'un arbre.	
*	Propriétés d'arbres binaires (relations entre le nombre de nœuds in-	S§5.5
	ternes et externes ou la hauteur).	
*	Parcours d'un arbre : préfixe/préordre, infixe/dans l'ordre, post-	S§5.6
	fixe/postordre, ordre de niveau.	
*	Arbre syntaxique. Conversions d'expressions arithmétiques : notations	S§4.3
	infixe, postfixe et préfixe.	

* Algorithmes récursifs sur les arbres : calcul de taille, hauteur ou profon-S§5.7

deur de sous-arbres.

File de priorité

Références

- ▶ Sedgewick §9.1–9.6
- Sedgewick & Wayne §2.4²⁰
- Notes sur la file de priorité : notes09-heap.pdf.

Sujets

- ★ Type abstrait de file de priorité²¹ min-tas/max-tas : opérations insert, deleteMin ou deleteMax. Implantations par tableau ou liste chaînée.
- * Arbre en ordre de tas²². Manipulation du tas : nager et couler (heapisation montante et descendante). Tas binaire²³, sa représentation dans un tableau.
- * heapify (établissement de l'ordre de tas dans un tableau); tri par tas²⁴, son temps de calcul et usage de mémoire.
- $\star\star$ Tas d-aire²⁵.

²⁰ http://algs4.cs.princeton.edu/24pg/

S§9.1,9.5

 21 W_(en):priority queue

22 W(fr):tas

S§9.2,9.3;SW§2.4

23 W(fr):tas binaire

S§9.4

 24 $W_{(fr)}$:tri par tas

25 W_(en):d-ary heap

Méthodes de tri

Références

- ▶ Sedgewick §6.1–6.4, §6.6, §6.9; chapitres 7, 8.
- \triangleright Sedgewick & Wayne §2.1²⁶, §2.2²⁷, §2.3²⁸.
- Notes sur les tris : notes10-tris.pdf.
- Notes sur le tri rapide : notes11-quicksort.pdf.

Sujets

- ★ Terminologie : tri interne et externe.
- * Tri par sélection²⁹ et tri par insertion³⁰.
- * Performances des tris élémentaires (pire cas, meilleur cas, cas moyen).
- * Fusion de tableaux.
- * Tri par fusion³¹ (descendant), sa performance.
- ★ Tri rapide³² : algorithme de base. Améliorations : partition par la médiane-de-trois, petits sous-fichiers.
- * Génération d'une permutation aléatoire
- * Performances du tri rapide (pire cas, meilleur cas, cas moyen)
- ** Preuve de la performance moyenne $O(n \log n)$ du tri rapide.
- \star Preuve de la borne inférieure $\lg(n!)$ sur le nombre de comparaisons au pire pour trier

- ²⁶ http://algs4.cs.princeton.edu/ 21elementary/
- ²⁷ http://algs4.cs.princeton.edu/22mergesort/
- ²⁸ http://algs4.cs.princeton.edu/23quicksort/

S§6.1

S§6.3,6.4; SW§2.1 $^{29}\,W_{(fr)}$:tri par sélection

30 W(fr):tri par insertion

S§6.6 S§8.1

S§8.3,8.4; SW§2.2

 31 W_(fr):tri fusion

 $^{32}\,W_{(fr)}$:tri rapide

S§7.1,7.4,7.5; SW§2.3 S§7.2,7.3; SW 2.3

SW§2.2

Arbres binaires de recherche

Références

▶ Sedgewick chapitres 12, §13.1, §13.3, §13.4

 \triangleright Sedgewick & Wayne §3.1³³, §3.2³⁴, §3.3³⁵.

Notes sur l'arbre binaire de recherche : notes12-abr.pdf.

▷ Notes sur l'arbre rouge et noir : notes13-rn.pdf.

33 http://algs4.cs.princeton.edu/

31elementary/

34 http://algs4.cs.princeton.edu/32bst/

35 http://algs4.cs.princeton.edu/33balanced

Sujets

* Type abstrait de la table de symboles.

* Recherche séquentielle et recherche binaire.

* Arbre binaire de recherche³⁶. Procédures fondamentales sur un ABR : recherche, insertion, suppression. Recherche de minimum ou maximum, successeur ou prédecesseur.

* Performance moyenne des opérations sur un ABR standard avec clés aléatoires.

* Notion d'un ABR équilibré. Maintenance d'équilibre : rotations simples et doubles.

* ABR rouge et noir³⁷. Définition par rang (hauteur noire) ou coloriage; équivalence des deux définitions. Coût des opérations dans le pire cas.

** Hauteur maximale d'un arbre rouge et noir.

★ Techniques de base sur les ABR rouges et noirs : promotion/rétrogradation, changement de couleur, rotation. Déroulement général d'une insertion ou suppression.

** Déroulement détaillé de l'insertion et de la suppression.

 \star Les arbres 2-3-4³⁸, et leur équivalence avec les arbres rouges et noirs. Techniques de base sur les arbres 2-3-4 : décalage et découpage, leur relation aux rotations et promotions.

S§12.6-12.9

S§12.1,12.2

S§12.3-12.5

 36 $W_{(fr)}$:ABR

S§13.1

³⁷ W_(fr):arbre bicolore

S§13.4

 38 W_(en):2-3-4 tree S§13.3

Algorithmes sur graphes

Références

- ▶ Sedgewick §3.7; Sedgewick & Wayne §4.1³⁹, §4.3⁴⁰
- ▶ Notes sur l'arbre couvrant minimal : notes14-acm.pdf.

Sujets

- * Représentation d'un graphe : matrice d'adjacence et listes d'adja-
- * Notion d'un arbre couvrant minimal⁴². Principe de base des algorithmes : la règle bleue. Logique générale des algorithmes de Prim⁴³ et de Kruskal⁴⁴, choix de structures de données.
- ** Analyse détaillé du temps de calcul des algorithmes. Choix de tas dans l'algorithme de Prim.

- 39 http://algs4.cs.princeton.edu/41graph/
- 40 http://algs4.cs.princeton.edu/43mst/

S§3.7;SW§4.1

- 41 W_(en):adjacency list
- 42 W(fr):ACM

SW§4.3

- $^{43}\,W_{(fr)}$:algorithme de Prim
- 44 W_(fr):algorithme de Kruskal









◀ français

Introduction

THIS DOCUMENT defines the skills and knowledge for the final examination in IFT2015, which is also the second part of the examen pré-doctoral in data

- **♦** Topics for a «B/A-» level are denoted by ★; ★★ denote somewhat more advanced topics for «A+/A» level.
- ★ The margin notes refer to the following books :
 - S Sedgewick, R. Algorithms in Java, Parts 1-4, 3rd edition (2003) SW Sedgewick, R. et K. Wayne. Algorithms, 4th edition (2011)
- * The class notes and links to Wikipedia articles are available on the webpage http://ift2015a16.wordpress.com/.
- * No documentation is allowed at the examen.



Principles of algorithm analysis

References

- ⊳ Sedgewick chapter 2
- Sedgewick & Wayne §1.4⁴⁵
- ▶ Notes on the foundations: notes01-recursion.pdf.
- Notes on algorithm analysis: notes04-analysis.pdf.

Topics

- * Basic principles: worst case, best case, average case.
- * Growth of common functions: constants, logarithms, polynomials, exponentials. Factorial (n!), Stirling's formula⁴⁶, Fibonacci numbers⁴⁷, harmonic numbers⁴⁸, iterated logarithm
- * Asymptotic notation 49 : definitions of big-Oh O(f), small-oh o(f), $\Theta(f)$, and $\Omega(f)$. Arithmetic expressions involving asymptotics, rules: O(f) + O(g), $O(f) \cdot O(g)$. Connections to lim

$$\lim_{n \to \infty} \frac{f(n)}{g(n)} = c > 0 \qquad \Rightarrow \qquad f(n) = O(g(n));$$

$$\lim_{n \to \infty} \frac{f(n)}{g(n)} = 0 \qquad \Leftrightarrow \qquad f(n) = o(g(n));$$

$$\lim_{n \to \infty} \frac{f(n)}{g(n)} = 1 \qquad \Leftrightarrow \qquad f(n) \sim g(n)$$

- * Using the definitions to prove f = O(g) or f = o(g).
- * Determination of space and time complexity for simple (iterative) algorithms, and for recursive algorithms (as a recursive expression).
- ★ Basic recurrences.

$$f(n) = f(n-1) + O(1) f(n) = O(n);$$

$$f(n) = f(n-1) + O(n) f(n) = O(n^2);$$

$$f(n) = f(n/2) + O(1) f(n) = O(\log n);$$

$$f(n) = f(n/2) + O(n) f(n) = O(n);$$

$$f(n) = 2f(n/2) + O(1) f(n) = O(n \log n);$$

- ** Proof by induction for asymptotic recurrences.
- * Notion of amortized cost.
- ** Proving amortized cost. Credit/debit method.
- * Experimental validation of running time

45 http://algs4.cs.princeton.edu/14analysis/

S§2.1,2.2,2.7 S§2.3 ⁴⁶ $n! \sim \sqrt{2\pi n} \left(\frac{n}{\rho}\right)^n$ $^{47} F_n = F_{n-1} + F_{n-2}$ $^{48}H_n = \sum_{i=1}^n 1/i = \ln n + \gamma + o(1)$ 49 W_(en):big-O notation S§2.4

S§2.5,2.6

Elementary structures and abstract data types

References

- ⊳ Sedgewick chapters 3 et 4
- ▷ Sedgewick & Wayne §1.1⁵⁰, §1.2⁵¹, §1.3⁵²
- Notes on linked lists: notes02-linkedlist.pdf.
- Notes on tables: notes03-tableaux.pdf.

50 http://algs4.cs.princeton.edu/11model/

52 http://algs4.cs.princeton.edu/13stacks/

Topics

- ★ Java building blocks.
- * Concept of an abstract data type, interface, implementation, client.
- * Abstract types for stacks, queues and generalized queues,
- ★ Linked lists⁵³. Variations: circular, doubly-linked lists. Sentinels⁵⁴ for head and/or tail. Manipulation of elements, insertion and deletion. List traversal.
- \star Arrays⁵⁵.
- * Implementations of stack and queue by tables or linked lists. Running time for standard operations in different implementations. Overflow/underflow.

S§3.1;SW§1.1

S§4.1;SW§1.2

S§4.2,4.7

53 W(en):linked list

S§3.3,3.4

 54 W_(en):sentinel

55 W_(en):array

S§3.2

S§4.4,4.5,4.7;SW§1.3

⁵¹ http://algs4.cs.princeton.edu/12oop/

Hash tables

Reference

- ▷ Sedgewick chapter 14.
- Sedgewick & Wayne §3.4⁵⁶, §3.5⁵⁷.
- Notes on hashing: notes06-hashing.pdf.

56 http://algs4.cs.princeton.edu/34hash/

57 http://algs4.cs.princeton.edu/ 35applications/

Topics

- ★ Basic notions for hashtables⁵⁸: load factor, collisions.
- * Hash functions : division and multiplication methods.
- * Collision resolution by separate chaining. Average-case performance with separate chaining as function of the load factor.
- * Open addressing: probe sequence. Search and insertion with open addressing. Lazy deletion, dynamic hashing. Linear probing, primary clustering. Double hashing.

S§14.1

 58 W_(en):hashtable

S§14.2

S§14.3-14.6

Union-find

References

- ▷ Sedgewick §1.2–1.3
- Sedgewick & Wayne §1.5⁵⁹
- ▷ Notes on Union-Find: notes07-unionfind.pdf.

59 http://algs4.cs.princeton.edu/15uf/

Topics

- * Connectivity problems, union-find operations.
- * Union-Find⁶⁰ data structure. Techniques: union-by-rank/union-bysize, path compression.
- ** Amortized cost per operation : $O(\alpha(m, n))$ for Union-Find with balanced trees and path compression; Ackermann⁶¹ function and its inverse

S§1.2

 60 $W_{\text{(en)}}$:union-find S§1.3;SW§1.5

61 W_(en):Ackermann function

Trees

D	C
кe	ferences

⊳ Sedgewick §4.3, §5.4–5.7

Notes on trees: notes08-trees.pdf.

Topics

*	Terminology for tree structures: k-ary tree, height, level, depth. Tree	S§5.4
	implementations.	
*	Mathematical properties of binary trees (relationships between number	S§5.5
	of internal and external nodes, height)	
*	Tree traversal: preorder, inorder, postorder, level-order.	S§5.6
*	Syntax tree. Conversion between arithmetic notations : infix, prefix	S§4.3
	and postfix.	
*	Recursions on trees: computing the size, height, or depth of subtrees.	S§5.7

.....

Priority queues

References

▶ Sedgewick §9.1–9.6

⊳ Sedgewick & Wayne §2.4⁶²

▶ Notes on priority queues: notes09-heap.pdf.

62 http://algs4.cs.princeton.edu/24pq/

Topics

- \star ADT for priority queue⁶³: operations insert, deleteMin or deleteMax. Implementations by table or linked list.
- * Heap⁶⁴ order for a tree. Heap manipulation : swim and sink. Binary heap⁶⁵, its representation in a table.
- * heapify (linear-time construction of heap order in a table); Heapsort⁶⁶, its running time and memory.
- $\star\star$ d-ary heap⁶⁷.

S§9.1,9.5

 63 $W_{\text{(en)}}\text{:priority}$ queue

64 W_(en):heap

S§9.2,9.3;SW§2.4

65 W_(en):binary heap

S§9.4

 $^{66}\ W_{\text{(en)}}\text{:}heapsort$

67 W_(en):d-ary heap

Sorting algorithms

References

- ▶ Sedgewick §6.1–6.4, §6.6, §6.9; chapters 7, 8.
- ▶ Sedgewick & Wayne §2.1⁶⁸, §2.2⁶⁹, §2.3⁷⁰.
- ▶ Notes on elementary sorting algorithms: notes10-tris.pdf.
- ▶ Notes on quicksort: notes11-quicksort.pdf.

Topics

- * Terminology: stable sort, internal and external sort.
- \star Insertion⁷¹ sort and selection⁷² sort.
- * Performance of elementary sorting algorithms (worst case, best case, average case).
- * Merging arrays.
- * Mergesort⁷³ (top-down), its performance.
- * Quicksort⁷⁴: basic algorithm. Improvements: pivoting by median-ofthree, small subarrays.
- * Performance of quicksort (worst case, best case, average case).
- * Generating a random permutation
- ** Proof of $O(n \log n)$ average running time for quicksort.
- \star Proof of the lower bound $\lg(n!)$ for the worst-case number of comparisons

- 68 http://algs4.cs.princeton.edu/ 21elementary/
- 69 http://algs4.cs.princeton.edu/22mergesort/
- 70 http://algs4.cs.princeton.edu/23quicksort/

S§6.1

 $^{71}\,W_{\text{(en)}} \\ \vdots \\ insertion \ sort$

 $^{72}\,W_{\text{(en)}}\text{:selection sort}$

S§6.3,6.4; SW §2.1

S§6.6 S§8.1

 $^{73}\;W_{(en)} ; merge\; sort$

S§8.3,8.4; SW§2.2

 $^{74}\ W_{(en)} \hbox{:} quick sort$

S§7.1,7.4,7.5; SW§2.3

S§7.2,7.3

SW§2.2

Binary search trees

Reference

- ▶ Sedgewick chapters 12, §13.1, §13.3 and §13.4
- \triangleright Sedgewick & Wayne §3.1⁷⁵, §3.2⁷⁶, §3.3⁷⁷.
- Notes on binary search trees: notes12-abr.pdf.
- Notes on red-black trees: notes13-rn.pdf.

- 75 http://algs4.cs.princeton.edu/ 31elementary/
- 76 http://algs4.cs.princeton.edu/32bst/
- 77 http://algs4.cs.princeton.edu/33balanced

Topics

- * Abstract data type of symbol table.
- ★ Sequential and binary search.
- * Binary search tree⁷⁸. Basic techniques: search, insertion, deletion. Searching for minimum or maximum, successor or predecessor.
- * Average performance of a standard BST with random keys.
- * Notion of a balanced BST. Maintaining the balance : simple and double rotations.
- * Red-black tree⁷⁹. Definition by rank (black height) or coloring; equivalence of the two definitions. Time complexity for operations in the worst-case.
- ** Maximum height of a red-black tree.
- * Basic techniques for red-black trees: promotion/demotion, recoloring, rotations. General outline of insertion and deletion.
- ** Detailed (case-by-case) steps in insertion and deletion.
- \star 2-3-4 tree⁸⁰, its equivalence with the red-black tree. Basic techniques with 2-3-4 trees: shifting and splitting, relationship with promotions and rotations in red-black tree.

S§12.1,12.2

S§12.3-12.5

 78 W_(en):BST

S§12.6-12.9

S§13.1

79 W_(en):red-black tree

S§13.4

80 W(en):2-3-4 tree

S§13.3

Graph algorithms

References

- ⊳ Sedgewick §3.7
- Sedgewick & Wayne §4.1⁸¹, §4.3⁸²
- ▷ Notes on minimum spanning tree and shortest path: notes14-acm.pdf.
- 81 http://algs4.cs.princeton.edu/41graph/
- 82 http://algs4.cs.princeton.edu/43mst/

Topics

- * Graph representations by adjacency matrix and adjacency lists⁸³.
- ★ Concept of a minimal spanning tree⁸⁴ (MST). Basic principles of MST algorithms: the blue rule (adding minimum-weight edge in a cut). General logic of Kruskal's⁸⁵ and Prim's⁸⁶ algorithms, choice of data structures. Basic justification for $O(n \log n)$ and $O(m \log n)$ running times (n nodes, m edges).
- ** Detailed analysis of running time for Kruskal's and Prim's algorithms. Choice of heap in Prim's algorithm.

S§3.7;SW§4.1

- $^{83}\ W_{(en)}\text{:adjacency list}$
- 84 W(en):MST

SW§4.3

- 85 W_(en):Kruskal's algorithm
- 86 $W_{\text{(en)}}\text{:}Prim\text{'s algorithm}$