

Università di Trieste

Laurea in ingegneria elettronica e informatica

Enrico Piccin - Corso di Algoritmi e Strutture dati - Prof. Andrea Sgarro

Anno Accademico 2021/2022 - 2 Marzo 2022

Indice

1 Introduzione	3
1.1 Architettura dei calcolatori	7
1.2 Diagramma di flusso	7
2 Ordinamento (sorting)	8
2.1 Bubble-Sort	8
2.2 Insertion-sort	9
2.3 Pseudocodice	9
2.3.1 Assegnazione	9
2.3.2 Condizione	10
2.3.3 Ciclo for	10
2.3.4 Ciclo while	10
3 Grafi	12
4 Algoritmi aritmetici	17
4.1 Algoritmo di Euclide	18
4.2 Notazione O grande	21
4.3 Funzioni di complessità	23
4.4 Classi di complessità	25
4.5 Merge-Sort	26
4.6 Master Theorem	30
4.7 Heap-sort	32
4.8 Quick-sort	48
4.9 Randomize Quick-Sort	51
4.10 Teorema del trasferimento	54
5 Ordinamento di strutture dati specifiche	66
5.1 Counting-Sort	66
5.2 Bucket-Sort	68
5.3 Riassunto degli algoritmi di ordinamento	73
6 Algoritmi sui grafi	78
6.1 Problema di Eulero	78
6.2 Algoritmo di Hamilton	79
6.3 Teorema delle strette di mano - Handshake theorem	79
6.4 Breadth-First-Search (BFS)	85
6.5 Depth-First-Search (DFS)	88
6.6 Teorema delle parentesi	94
6.7 Teorema del cammin bianco	95
6.8 Teorema di classificazione degli archi	97
6.9 Classificazione	100

6.10 Ordinamento topologico	100
6.11 Algoritmo di Dijkstra	101

1 Introduzione

Algoritmo è una parola molto antica, non connessa all'utilizzo e all'invenzione del calcolatore. Alla base della teoria della computazione si pone il **Liber abaci** (tradotto "Libro della computazione"), scritto nel 1200 da Leonardo Bonacci, in contatto con la popolazione araba, in quanto mercante; egli è venuto a conoscenza della **numerazione araba**, introducendola in Occidente e spiegandola dettagliatamente all'interno del **Liber abaci**.

La numerazione araba è uno straordinario passo in avanti nella scienza, in quanto con essa viene introdotto il concetto di **notazione posizionale**, così come l'importanza del numero 0: i numeri non servono solamente per contare, come si pensava in precedenza, e per questo rinnegando il numero 0.

A Firenze, sempre negli stessi anni, ci fu una **protesta sindacale** contro l'innovazione tecnologica, contro questa nuova scoperta, facendo pressione affinché il governo abolisse il nuovo sistema di numerazione, in quanto avrebbe fatto perdere il posto di lavoro a tutti coloro che prima eseguivano difficili calcoli con la numerazione romana: tuttavia, tale proteste, com'è noto, possono rallentare il progresso, ma mai arrestarlo.

Leonardo Bonacci, nei suoi viaggi in Oriente, venne a conoscenza del **Liber abaci** di Al-Gorasmy, proveniente dalla Coresmia, ma che parlava persiano, da cui poi sarebbe stato tratto il nome **Algoritmo**, che letteralmente significa **procedimento di calcolo**.

ALGORITMO

Algoritmo significa letteralmente **procedimento di calcolo**. Tuttavia, bisogna chiarire che un algoritmo è un procedimento di calcolo non necessariamente numerico, ma molto più generale, che va ben al di là dei numeri.

Un altro importante elemento che contraddistingue l'algoritmo è la **meccanicità**, ovvero la sua esecuzione può essere affidata ad una macchina: ciò significa che un algoritmo non deve necessariamente essere meccanizzato, ma deve essere **meccanizzabile**; in altre parole, l'esecuzione (bada bene, l'esecuzione e non la sua ideazione) dell'algoritmo è completamente **stupida**.

Esempio 1: L'algoritmo della moltiplicazione è molto chiaro e semplice: basta solamente accedere ad una **base di dati** in cui sono memorizzati i prodotti elementari fra numeri molto piccoli, e quindi molto più semplici da trattare.

Una volta eseguite le operazioni di moltiplicazione tramite quanto esposto in precedenza, è necessario eseguire delle operazioni di addizione, che era necessario aver precedentemente memorizzato. Ecco che quello che si è appena eseguito è un **procedimento di calcolo**.

Esempio 2: L'algoritmo della divisione fa sempre uso di una base di dati, nella quale devono essere memorizzati i risultati del prodotto del divisore con tutti i numeri decimali da 0 a 9 e confrontare ciascun prodotto con il termine da dividere per ottenere il quoziente.

Alternativamente, si sarebbe potuto creare un ciclo da 0 a 9 in cui per ogni indice si sarebbe dovuto verificare se questo fosse il fattore moltiplicativo corretto per ottenere la quantità giusta da sottrarre.

Osservazione: In ciascuno di tali esempi è essenziale la meccanicità del processo esecutivo, che appare evidente.

Quando si rappresentano delle quantità e, a maggior ragione, quando si effettuano dei calcoli, è fondamentale fissare una base di rappresentazione, da cui poi dipendono le cifre che si possono impiegare per la rappresentazione stessa.

La notazione posizionale permette anche di comprendere la rappresentazione di qualsiasi quantità con qualsiasi base, effettuando anche delle conversioni di base a seconda della maggiore o minore convenienza di rappresentazione.

Per esempio, volendo convertire una quantità rappresentata in base $\mathcal{B} = 7$ in una base $\mathcal{C} = 10$ si deve procedere come segue

$$(5203)_7 = 5 \cdot 7^3 + 2 \cdot 7^2 + 0 \cdot 7^1 + 3 \cdot 7^0 = 1715 + 98 + 0 + 3 = (1816)_{10}$$

Ovviamente le basi di rappresentazione sono almeno binarie, in quanto la **base unaria** non può, per ovvie ragioni, rappresentare alcuna quantità se non quella unica che viene permessa dalla base scelta, ossia lo 0.

Tuttavia, il processo inverso, atto a passare dalla rappresentazione di una quantità in base 10 ad una in base 3, non risulta essere così immediato.

Per cercare un algoritmo che permette di effettuare tale conversione, si effettua un primo **passaggio contointuitivo** (che suggerisce, tuttavia, il corretto processo esecutivo), che prevede di rappresentare una quantità in base 10 in una quantità ancora in base 10, tramite un processo di divisioni successive. Si consideri, a tal proposito

$$(3412)_{10}$$

e si divida progressivamente tale numero per 10, come segue

3412	10
341	2
34	1
3	4
0	3

Leggendo, ora, i resti, al contrario si ottiene il numero cercato all'inizio. Se ora si prova a considerare un'altra base, come 3, l'operazione porta ad un risultato analogo

3412	3
1137	1
379	0
126	1
42	0
14	0
4	2
1	1
0	1

Per cui si è ottenuto

$$(3412)_{10} = (11200101)_3$$

Scegliendo la base 2 si ottiene, per esempio

241	2
120	1
60	0
30	0
15	0
7	1
3	1
1	1
0	1

Per cui si è ottenuto

$$(241)_{10} = (11110001)_2$$

Ovviamente la lunghezza di rappresentazione in base 2 prevede un numero di cifre pari a circa il triplo di quelle impiegate per rappresentare la medesima quantità in base 10, proprio perché

$$\log_2(10) \cong 3.3$$

Per passare da base 10 a base 100, le operazioni sono molto semplici

$$(375712)_{10} = [(37)(57)(12)]_{100}$$

usando come simboli

$$(00), (01), \dots, (75), \dots, (99)$$

Si consideri, ora la base 8 e si scriva un numero binario in base ottale:

$$(010101010)_2 = [(010)(101)(010)]_8 = (252)_8$$

Ancora una volta, le cifre impiegate per la rappresentazione sono state ridotte ad un terzo, sempre perché

$$\log_2(8) = 3$$

E se ora si volesse impiegare la base 16 si otterrebbe:

$$(010101010)_2 = [(1010)(1010)]_{16} = (\text{AA})_{16}$$

Osservazione: Si consideri una lunghezza $l = 5$. Allora usando 5 cifre, non tutte nulle, in base 10, i numeri n che si possono rappresentare sono

$$10000 \leq n \leq 99999 \quad \equiv \quad 10^4 \leq n < 10^5 \quad \equiv \quad 10^{l-1} \leq n < 10^l$$

Da ciò si può estrapolare un risultato importante

$$\log_{10}(10^{l-1}) \leq \log_{10}(n) < \log_{10}(10^l) \quad \equiv \quad l - 1 \leq \log_{10}(n) < l$$

che è una relazione esatta. Tuttavia, approssimativamente, si può scrivere che

$$l_{10}(n) \cong \log_{10}(n)$$

3 Marzo 2022

Com'è noto, il matematico indiano **Ramanujan** ha affermato che la matematica esatta non rappresenta una base solida per la realtà, mentre la matematica vera è fatta di approssimazioni.

Hardy scoprì quanto fosse importante lo studio di **Ramanujan** e insieme a lui portò avanti la teoria dei numeri, una teoria **asintotica** che, come lui stesso affermava, non può essere esatta, ma fatta di approssimazioni.

Se, per esempio, si considera una quantità scritta in base 5, quale $n = (412)_5$

$$(412)_5 = 4 \cdot 5^2 + 1 \cdot 5^1 + 2 \cdot 5^0 = (107)_{10}$$

Se, ora, si fissa una lunghezza $l = 4$, una lunghezza rigida, senza considerare zeri in testa, si può capire che con 4 cifre si possono rappresentare, in base 5 numeri n nell'intervallo

$$1000 \leq n < 10000$$

ovvero tale per cui

$$5^{l-1} \leq n < 5^l$$

e ciò funziona con qualsiasi base, per cui, in generale, fissata una lunghezza l e una base \mathcal{B} si ha che le quantità che possono essere rappresentate con l cifre, non tutte uguali a 0 è

$$\mathcal{B}^{l-1} \leq n < \mathcal{B}^l$$

Traducendo tale risultato tramite il logaritmo in base \mathcal{B} , sfruttando la crescenza in senso stretto della funzione logaritmica, si ottiene, equivalentemente

$$l - 1 \leq \log_{\mathcal{B}}(n) < l$$

in cui, ovviamente,

$$\log_{\mathcal{B}}(n) < l \leq \log_{\mathcal{B}}(n) + 1$$

che si può scrivere che

$$l_{\mathcal{B}}(n) \cong \log_{\mathcal{B}}(n)$$

Per cui l'**errore massimo** che si può commettere è di 1 cifra in base \mathcal{B} , nel caso peggiore, ma sarà sempre un po' maggiore del $\log_{\mathcal{B}}(n)$, per cui il logaritmo è una **sottostima della lunghezza**. Tuttavia, nello spirito di Hardy, sarà utile anche scrivere che la lunghezza binaria di n è circa uguale al logaritmo binario di n , ovvero

$$\log_{\mathcal{B}}(n) \cong \log_2(n)$$

in cui si può interpretare il $\log_{\mathcal{B}}(n)$ come una **lunghezza analogica**, mentre $l_{\mathcal{B}}(n)$ è una **lunghezza digitale**, in quanto **intera**, con precisione alla cifra (senza nulla in mezzo): in molti casi sarà più utile la lunghezza analogica di quella digitale, in quanto molto più precisa.

Grazie a questa formula è possibile capire facilmente come si alterano le lunghezze quando si effettua un cambiamento di base. Per esempio, si può osservare che

$$\log_2(10) \cong 3.38$$

per cui la lunghezza in base 2 è circa tre volte la lunghezza in base 10.

Esempio: Per trasformare un numero da base 10 in base 5 si deve procedere per divisioni successive per 5, considerando i resti (per questo si parla di *divisione intera*). Per esempio si ha che

$$10 \div 3 = 3 \text{ con resto di } 1$$

in cui, ovviamente, il resto r può essere

$$0 \leq r < D$$

con D divisore. Convertendo 32 da base 10 a base 5 ci si aspetta di ottenere un resto $0 \leq r \leq 4$.

1.1 Architettura dei calcolatori

Il calcolatore, naturalmente, si basa sulla logica binaria, ovvero opera impiegando la rappresentazione in base 2.

Il metodo più utilizzato per rappresentare caratteri diversi da quelli binari, tramite una codifica binaria, è il metodo ASCII (dall'inglese, American Standard Code For Information Interchange). Naturalmente, siccome la codifica tramite ASCII fa uso di soli 7 bit (sarebbero 8, ma un bit è riservato alla parità, per la rilevazione degli errori), il numero di n -uple binarie che si possono ottenere è 2^7 , un numero certamente irrisorio per la rappresentazione di tutti i caratteri alfanumerici necessari per la comunicazione multilingua.

In generale, fissata una lunghezza n , il numero di n -uple binarie distinte è, ovviamente, 2^n , che rappresenta una crescita esponenziale, praticamente infinita, anche se, ovviamente, in teoria sono un numero ben limitato.

Esempio: Si consideri una macchina calcolatrice che considera delle istruzioni di 9 bit come di seguito esposto:

NOME ISTRUZIONE	ISTRUZIONE
ADD	010 $\times \times \times \times \times$
PUNCH	100 $\times \times \times \times \times$

Tabella 1: Tabella di istruzioni operative per un calcolatore

Naturalmente, tale linguaggio è **Assembly**, ovvero un linguaggio molto simile al linguaggio macchina, che risulta particolarmente complesso da impiegare per lo sviluppo di software.

1.2 Diagramma di flusso

Si consideri il seguente **flowchart**, o **diagramma di flusso**:

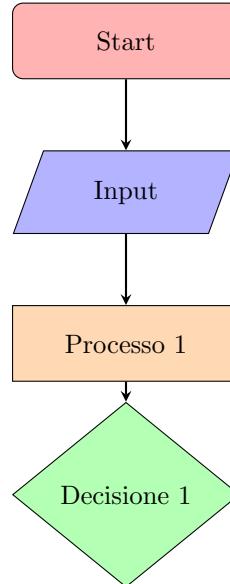


Figura 1: Diagramma di flusso

Tuttavia, tale tecnica di progettazione algoritmica è oramai superata, lasciando il posto allo **pseudocodice**, ossia un linguaggio di definizione delle istruzioni slegato da qualsiasi specifico linguaggio di programmazione di riferimento, che permette di esporre una serie di istruzioni esecutive molto simili a quelle di un programma vero e proprio.

2 Ordinamento (sorting)

Si espongono, di seguito, i principi di algoritmica dei più importanti algoritmi di ordinamento. Ciascuno di tali algoritmi prevede di effettuare l'ordinamento di n numeri forniti come input, in modo debolmente crescente, in caso di uguaglianza.

2.1 Bubble-Sort

Si espone di seguito l'algoritmo di ordinamento **bubble-sort** impiegando lo *pseudocodice*:

Algorithm 1 Bubble-sort

- 1: **do** the following $n - 1$ times
 - 2: *point* to the 1st element
 - 3: **do** the following $n - 1$ times
 - 4: *compare* with next
 - 5: **if** wrong order *exchange*
 - 6: *point* to the next
-

Naturalmente tale algoritmo è corretto e lo si può verificare immediatamente, considerando, per esempio, i seguenti 5 elementi, così ordinati:

2 3 1 4 5

Ovviamente il procedimento ci porta ad eseguire l'algoritmo 4 volte. Nella prima iterazione si ottiene

2 1 3 4 5

la seconda iterazione, invece, porta ad ottenere

1 2 3 4 5

mentre le ultime due iterazioni sono superflue. Si capisce facilmente che tale algoritmo non risulta essere pienamente efficiente, in quanto alcune iterazioni potrebbero essere evitate, tramite un **flag**, per esempio. Allo stato attuale, il numero delle iterazioni da eseguire è

$$\#\text{iterazioni} = (n - 1) \cdot (n - 1)$$

Se, invece, si facesse in modo di evitare alcune iterazioni si avrebbe un numero di iterazioni

$$(n - 1) \leq \#\text{iterazioni} \leq (n - 1)^2$$

considerando $n - 1 \cong n$, e quindi $(n - 1)^2 \cong n^2$ si può dire che la complessità del bubble-sort è **quadratica**, in quanto il numero delle iterazioni è n^2 .

4 Marzo 2022

L'algoritmo bubble-sort non viene utilizzato, ad oggi, così come non si impiega lo pseudocodice in *pseudo-english* (da leggere psude-english). Esso è funzionale, ma non efficiente, in quanto la sua complessità è n^2 .

Ecco che per definire un algoritmo di ordinamento non è necessaria solamente la sua funzionalità, ma anche l'efficienza.

2.2 Insertion-sort

L'insertion-sort è un algoritmo di ordinamento che prevede di considerare ciascuna quantità da ordinare ad una ad una e di effettuare un confronto solo quando ci sono dei cambiamenti.

La prima quantità è ovviamente già in ordine con se stessa. Se la seconda è più piccola della prima, si effettua uno scambio, per cui ora i primi due numeri sono ordinati. Si considera, ora, il terzo numero e se questo è più piccolo del secondo si effettua uno scambio e un nuovo confronto tra la seconda e la prima e così via.

Pertanto si effettuano tutti i confronti solamente quando si ha uno scambio delle quantità: questo comporta che il minimo numero di iterazioni è $n - 1$, se n è il numero delle quantità da ordinare. Se, invece, tutte le quantità sono in disordine si effettua un numero di iterazioni pari

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

una formula molto semplice che Gauss determinò come segue, ovverosia scrivendo la somma dei numeri da 1 a k in ordine crescente e poi decrescente, come mostrato di seguito:

$$\begin{array}{cccccccccc} 1 & + & 2 & + & 3 & + & \dots & + & k - 1 & + & k \\ k & + & k - 1 & + & k - 2 & + & \dots & + & 2 & + & 1 \end{array}$$

essendo k numeri in ambedue le righe, sommando i due termini corrispondenti, uno sotto l'altro, si ottiene sempre $k + 1$ che, sommato per k volte produce $k \cdot (k + 1)$. Tuttavia, dal momento che tale quantità è il doppio di quella richiesta si ottiene

$$\frac{k \cdot (k + 1)}{2}$$

Pertanto si ha che il numero di iterazioni dell'algoritmo *insertion-sort* è

$$n \cong n - 1 \leq \#\text{iterazioni} \leq \frac{n \cdot (n - 1)}{2} \cong n^2$$

che, in maniera approssimata è

$$n \leq \#\text{iterazioni} \leq n^2$$

pertanto, nel caso migliore, la **complessità è lineare**, mentre nel caso peggiore, la **complessità è quadratica**.

2.3 Pseudocodice

Per la scrittura dello pseudocodice si devono impiegare delle notazioni e dei simboli ben specifici, che di seguito vengono riportati.

2.3.1 Assegnazione

L'**assegnazione** viene indicata con il simbolo $=$ (oppure $:=$ o \leftarrow), anche se l'assegnazione non è un'uguaglianza. Per esempio, la notazione

$$A = 3$$

significa che nella cella di memoria A viene inserito il valore 3. Analogamente, se si scrive

$$A = A + 1$$

significa che il valore presente nella cella di memoria A viene incrementato di 1 unità.

2.3.2 Condizione

La specifica della **condizione** avviene tramite l'istruzione **if**, secondo la notazione seguente:

```
if C then
    istruzioni
else
    istruzioni
```

2.3.3 Ciclo for

Il **ciclo for** è un'istruzione di ciclo in cui vengono indicate specificatamente le iterazioni che devono essere eseguite, secondo la notazione seguente

```
for  $i = 0$  to  $n$  do
```

2.3.4 Ciclo while

Il **ciclo while** è un'istruzione di ciclo in cui si effettuano le istruzioni fintantoché la condizione specificata è vera

```
while  $C$  do
```

Si consideri lo pseudocodice dell'algoritmo **INSERTION-SORT(A)**, esposto di seguito, dove A sta ad indicare **array**, ovvero un record di $\text{length}[A] = n$ valori da ordinare, già forniti in input. Di seguito si espone lo pseudocodice, in cui si parte da 1 come posizione iniziale dell'array:

Algorithm 2 Insertion-sort

```
1: for  $j = 2$  to  $\text{length}[A] = n$ 
2:   do  $key = A[j]$ 
3:   ...
4:    $i = j - 1$ 
5:   while  $i > 0 \wedge A[i] > key$ 
6:     do  $A[i + 1] = A[i]$ 
7:        $i = i - 1$ 
8:        $A[i + 1] = key$ 
```

Tale codice è concluso e il suo funzionamento può essere facilmente verificato come segue, considerando l'array A di lunghezza $\text{length}[A] = 6$:

5	2	4	6	1	3
---	---	---	---	---	---

Partendo con $j = 2$ si fissa $key = A[j] = 2$ e imponendo $i = 1$, si entra all'interno del ciclo *while*, in quanto $i > 0$ e $A[i] > key$ e si effettua l'istruzione $A[i + 1] = A[i]$ e $A[i + 1] = key$, trovandosi nella configurazione seguente

2	5	4	6	1	3
---	---	---	---	---	---

Terminato il ciclo while e il ciclo for si procede con $j = 3$ si fissa $key = A[j] = 4$ e imponendo $i = 2$, si entra all'interno del ciclo *while*, in quanto $i > 0$ e $A[i] > key$ e si effettua l'istruzione $A[i + 1] = A[i]$ e $A[i + 1] = key$, trovandosi nella configurazione seguente

2	4	5	6	1	3
---	---	---	---	---	---

Adesso, terminato il ciclo while e for, si considera $j = 4$, specificando $key = A[j] = 6$ e $i = 3$. In questo caso, tuttavia, non si entra nel ciclo while, in quanto $i > 0$, ma $A[i] < key$. Si procede direttamente con $j = 5$, $key = A[j] = 1$ e $i = 4$ e si entra nel ciclo while, compiendo tutte le iterazioni

2	4	5	1	6	3
2	4	1	5	6	3
2	1	4	5	6	3
1	2	4	5	6	3

e così via fino ad arrivare all'ordinamento finale

1	2	3	4	5	6
---	---	---	---	---	---

Ecco che, come si può vedere, tale algoritmo ha una complessità che, nel caso migliore, è **lineare** (n) e nel caso peggiore è **quadratica** (n^2).

Mentre si chiama **complessità tipica** la **complessità media**, ovvero la complessità dell'algoritmo nel caso intermedio. In questo caso la complessità tipica è n^2 , che può essere calcolata, in maniera non propriamente corretta, come segue

$$\frac{n + n^2}{2} = n^2$$

Esiste, infine, anche una **complessità empirica**, basata sull'uso pratico dell'algoritmo: in particolare, l'algoritmo insertion-sort funziona particolarmente bene quando il **numero degli elementi da ordinare è ridicolmente basso**, il che potrebbe essere un controsenso; tuttavia, potrebbe essere particolarmente utile ricorrere all'ordinamento di pochi numeri all'interno di una procedura particolarmente complessa: ecco, allora, che l'utilizzo di insertion-sort diviene conveniente (cosa che non accade per bubble-sort).

Osservazione: È importante osservare che l'algoritmo di insertion-sort è un **algoritmo di ordinamento in loco**, ovvero tale per cui non si impiega un altro array per l'ordinamento, ma tutte le operazioni si effettuano sullo stesso array di partenza.

Osservazione: Quando si parla di algoritmica, non è possibile parlare di **completezza** senza parlare di **complessità**.

3 Grafi

Il **grafo** è una **struttura finita**. Gli elementi costituitivi di un grafo sono i **vertici** (o **nodi**) e gli **archi** (o **lati**) (dall'inglese *arcs* o *edges*). Per indicare i vertici si impiega la lettera v , mentre per indicare gli archi si usa la lettera ξ .

Un arco collega due vertici distinti che, per il momento, non è orientato, non rappresenta una freccia, in quanto si parla di **grafi semplici**, come illustrato di seguito:

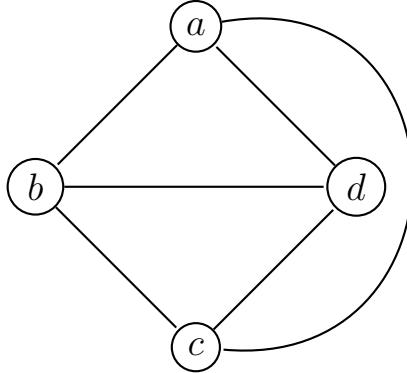


Figura 2: Esempio di grafo semplice

Il calcolo della copertura dei vertici prende il nome di **vertex cover** e prevede di definire il numero minimo di vertici essenziali per coprire tutti gli archi. L'ottimizzazione del grafo, in questo caso, prevede di determinare la copertura minima.

Si supponga di avere a disposizione k vertici (che si indica come $|v| = k$, in cui $|v|$ rappresenta la **cardinalità** dell'insieme dei vertici). Naturalmente, la copertura minima pari a 0 si ha quando i vertici del grafo sono tutti scollegati, ovvero non ci sono archi.

Il numero di archi in un grafo completo è pari a

$$\frac{|v| \cdot (|v| - 1)}{2}$$

per cui si potrebbe, in questo caso limite, affermare che la copertura minima sia k , ma se si elimina un nodo ancora la copertura sussiste, in quanto su ogni arco vi sarà sempre almeno un nodo coperto. Quindi si può affermare che la *vertex cover*, nel caso generale è compresa tra 0 e $k - 1$, con k numero dei vertici.

9 Marzo 2022

Il problema del **vertex cover**, ovvero di “ricoprimento dei vertici”, è un problema che riguarda la teoria dei grafi.

Gli elementi costitutivi di un grafo sono i **vertici** (o **nodi**, molto più raramente chiamati *punti*) e gli **archi** (dall’inglese *edges*, traducibili in **spigoli** o, più impropriamente, in *lati*), per il momento non orientati, che collegano due vertici, per il momento necessariamente distinti.

La notazione per indicare vertici e archi è la seguente

- L’insieme dei vertici si denota con v
- L’insieme degli archi si denota con ξ

Naturalmente, sussiste la possibilità che in un grafo tutti i vertici siano sconnessi ed **isolati**, ovvero il numero degli archi sia nullo, per cui si ottiene che $|\xi| = 0$.

Analogamente, volendo collegare tutti i nodi con un arco (ottenendo un **grafo completo**), si procede come segue: partendo da un primo vertice se ne collega un secondo, necessariamente distinto; ma non volendo considerare ogni arco due volte, si divide per due, ottenendo

$$\frac{|v| \cdot (|v| - 1)}{2} \cong |v|^2$$

da cui si evince che il numero degli archi, in un grafo, è compreso tra

$$0 \leq |\xi| \leq \frac{|v| \cdot (|v| - 1)}{2}$$

considerando $|v|$ la cardinalità, ossia il numero dei vertici considerati che costituiscono il grafo.

Il problema di **vertex cover** è un problema di ottimizzazione: ridurre il numero minimo di vertici tale per cui nel grafo ad ogni arco deve essere collegato almeno un vertice coperto. Per esempio, in un grafo dove ogni nodo è isolato, il numero minimo dei vertici da coprire è 0, in quanto non ci sono archi. Nel caso di un **grafo completo**, come quello esposto di seguito

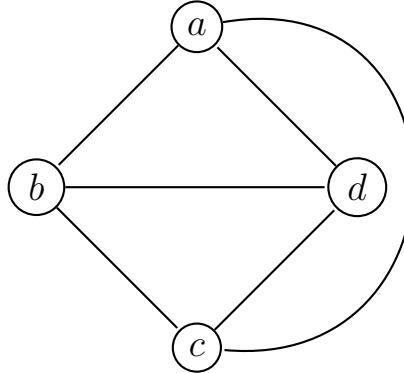


Figura 3: Esempio di grafo semplice completo

la copertura minima non è, come contrariamente si potrebbe pensare all’inizio, pari a $|v|$, in quanto eliminando uno qualsiasi dei nodi, ancora ad ogni arco sarà collegato almeno un nodo comperto. Pertanto si ha che il numero $\#nodi$ dei nodi che si dovranno coprire al fine di risolvere il problema del vertex cover in un grafo avente $|v|$ vertici sarà sempre compreso tra

$$0 \leq \#nodi \leq |v| - 1$$

Per la risoluzione meccanica del **vertex cover** vi sono due algoritmi, di cui solo il secondo realmente efficace. Si consideri, a titolo esemplificativo, il grafo seguente

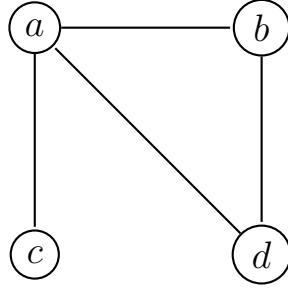


Figura 4: Esempio di grafo semplice

in cui, naturalmente, per coprire tutti gli archi sarà sufficiente considerare il vertice a e il vertice d (oppure il nodo b), ottenendo, come possibile copertura

a	b	c	d
1	0	0	1

in cui con $\boxed{1}$ si rappresenta la copertura del rispettivo vertice. Ecco che questa è una k -upla binaria di **peso** $w = 1 + 1 = 2$, in cui $k = 4$: usando tale notazione, quindi, un procedimento meccanico, atto a verificare la corretta copertura prevede di considerare tutti gli archi e verificare per ciascuno di essi che almeno ad un vertice collegato dall'arco in questione corrisponda un 1; se un arco è collegato a due vertici cui corrisponde 0, la copertura è scorretta, ovviamente.

Tuttavia, la k -upla 1001 è anche una codifica binaria su 4 bit del numero $(9)_{10}$, che suggerisce la procedura di controllo seguente, definita a partire da k numero di vertici

Algorithm 3 Vertex-cover

```

1: for  $i = 0$  to  $2^k$ 
2:    $i \rightarrow$  binario
3:    $check(i)$ 

```

ove $check(i)$ è una procedura che svolge il compito precedentemente esposto: presa in ingresso una k -upla binaria, cui si fa corrispondere la copertura di rispettivi vertici, verifica per ciascun arco che ad esso sia collegato un vertice effettivamente coperto, ovvero cui corrisponde un 1 nella k -upla binaria considerata; alla fine della procedura si ottiene la copertura di peso minore, ovvero quella con meno 1 e quindi che prevede meno vertici coperti.

Tale algoritmo non risulta propriamente efficiente; pertanto, al fine di eliminare alcune iterazioni, si potrebbe pensare di partire con il peso w ed effettuare il medesimo $check$ per tutte n -uple di peso w , come mostrato di seguito:

Algorithm 4 Vertex-cover

```

1: for  $w = 0$  to  $k - 1$ 
2:   for all

```

che è un procedimento più razionale, in quanto si controlla progressivamente se sono sufficienti un numero sempre maggiore di vertici da coprire (basta 1 vertice? Bastano 2 e così via...) e si esce dal ciclo quando si trova il primo, in quanto quella copertura sarà certamente la minima.

Da notare che si cicla fino a $k - 1$ e non fino a k , in quanto è noto dalla teoria che il numero di vertici che si andranno a coprire per risolvere un qualsiasi problema di vertex-cover è sempre compreso tra 0 e $k - 1$, con k numero di vertici: ma l'unica n -upla con peso massimo è quella con tutti 1, la quale costituisce una sola configurazione tra le 2^k possibili. Pertanto, nel caso peggiore, si potrebbe procedere ad effettuare un numero di iterazioni pari a $2^k - 1$, che non è molto dissimile dal caso precedente, in cui si facevano inevitabilmente 2^k iterazioni.

Questo algoritmo, pertanto, pur essendo corretto, è ispirato al meccanismo dell'**exhaustive search** (dall'inglese, ricerca esauriente), che prevede di controllare tutti gli archi al fine di verificarne la

corretta copertura.

Il motivo per cui tale algoritmo è inutilizzabile è che presenta un numero di **iterazioni esponenziale** e, conseguentemente, una **complessità esponenziale**, la quale è intollerabile, dal momento che il numero delle iterazioni cresce esponenzialmente al variare dell'input.

Di seguito si espone, invece, un nuovo algoritmo che risolve il problema del **vertex cover**; alla base di tale algoritmo si pone la seguente idea: si considerino dapprima due nodi connessi da un arco e si eliminino tutti gli archi che incidono sul primo e sul secondo vertice e si proceda a considerare un nuovo arco che insiste su due nodi ancora non considerati e si eliminino tutti gli altri archi che incidono sui nodi stessi e così via, fino ad esaurire tutti gli archi a disposizione, tale che alla fine della procedura si ottiene un ricoprimento vero e proprio.

Come di consueto, nello pseudocodice esposto di seguito, l'input non viene specificato a priori, ma il numero delle iterazioni per specificare l'input è noto, ossia è pari a $\cong |v| + |\xi|$.

Lo pseudocodice è il seguente:

Algorithm 5 Vertex-cover

```

1:  $C \leftarrow \emptyset$ 
2:  $E' \leftarrow \xi(\mathcal{G})$ 
3: while  $E' \neq \emptyset$ 
4:   do ...  $(u, v)$  in  $E'$ 
5:      $C \leftarrow C \cup \{u, v\}$ 
6:     delete incidenti
7: return  $C$ 
```

In cui C è un contenitore, inizialmente vuoto, all'interno del quale successivamente andranno inseriti i vertici necessari per la vertex cover. Invece, E' è un contenitore, inizialmente pieno, in quanto contiene tutti gli archi che dovranno essere esaminati/scartati. Dopodiché, fintantoché non ci sono più archi da analizzare, si considera un primo arco, designato con la notazione (u, v) , i cui due estremi, appunto u e v andranno ad essere inseriti all'interno di C , mentre verranno eliminati da E' tutti gli archi incidenti in u o in v , finché non si avranno più archi a disposizione.

Il numero di iterazioni in cui tale algoritmo si impegna è, approssimativamente, pari a $|v| + |\xi|$, ovvero si ha una **complessità lineare**: questa è un'ottima notizia, in quanto significa che tale algoritmo è velocissimo; l'unico problema è che esso è scorretto.

Si consideri, a titolo di esempio, il seguente grafo semplice:

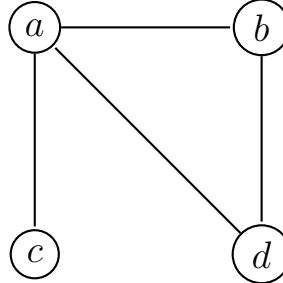


Figura 5: Esempio di scorrettezza dell'algoritmo considerato

In questo caso, l'algoritmo, considerando per primo l'arco $a - b$, elimina tutti gli altri archi e produce un risultato corretto.

Tuttavia, se il primo arco considerato dall'algoritmo fosse $a - c$, l'algoritmo è costretto a considerare anche l'arco $b - d$, non producendo un risultato corretto.

Tale algoritmo prende il nome di **algoritmo approssimato**, in quanto non sempre produce un risultato corretto, ma mai disastroso.

Tuttavia, non esiste un algoritmo che sia corretto e anche accettabile dal punto di vista del numero delle iterazioni e quindi della complessità: pertanto ci si deve accontentare dell'algoritmo approssimato appena esposto.

Per capire il range di risposta di tal algoritmo, si assuma che il numero di nodi ottimali necessari alla copertura sia $|\mathcal{O}|$ e il numero dei nodi effettivamente ottenuti dalla procedura algoritmica sia

$|\mathcal{P}|$, legati dalla seguente relazione, dimostrata con l'esempio precedente:

$$|\mathcal{O}| \leq |\mathcal{P}|$$

Naturalmente, per quanto visto con l'esempio precedente, tale disuguaglianza può anche essere stretta. Ovviamente, se ora si considerano gli archi privilegiati dall'algoritmo esposto, denotati con $|\mathcal{A}|$, ovverosia l'insieme degli archi che l'algoritmo ha via via considerato, esso, naturalmente, presenta la seguente cardinalità

$$|\mathcal{A}| = \frac{|\mathcal{P}|}{2}$$

in quanto su ogni arco vi sono due nodi (u e v nello pseudocodice), per cui basta considerarne la metà. Ora, il numero di nodi ottimale $|\mathcal{O}|$ deve essere necessariamente almeno uguale al numero di archi privilegiati $|\mathcal{A}|$, dal momento che ciascuno di tali archi deve insistere almeno su un vertice coperto. Pertanto $|\mathcal{O}|$ e $|\mathcal{A}|$ sono legati dalla seguente relazione

$$|\mathcal{O}| \geq |\mathcal{A}|$$

per cui si ottiene che

$$\frac{|\mathcal{P}|}{2} \leq |\mathcal{O}| \leq |\mathcal{P}|$$

4 Algoritmi aritmetici

L'**algoritmo di Euclide** permette di calcolare il **Massimo Comune Divisore (M.C.D.)** tra due numeri, ma non procedendo alla fattorizzazione in fattori primi tra le due quantità considerate. Infatti, normalmente, per determinare l'M.C.D. tra due quantità è necessario procedere alla scomposizione delle due in fattori primi, come mostrato di seguito per i numeri 12 e 9:

$$\begin{aligned}12 &= 3 \cdot 2 \cdot 2 \\9 &= 3 \cdot 3\end{aligned}$$

da cui si evince che

$$\text{MCD}(12, 9) = (12, 9) = 3$$

Tuttavia, non è possibile procedere attraverso la fattorizzazione per la risoluzione di tale problema, in quanto gli algoritmi per la fattorizzazione sono estremamente lenti. L'algoritmo di Euclide, invece, risolve tale problematica in maniera corretta, veloce ed efficiente, basandosi sul meccanismo della **ricorsività** e delle divisioni intere. Infatti, com'è noto, la divisione può essere di due tipologie

1. Divisione esatta: $10 \div 3 = 3, \bar{3}$
2. Divisione intera: $10 \div 3 = 3$ con resto 1

10 Marzo 2022

Naturalmente, un algoritmo non può essere applicato concretamente se ha una crescita esponenziale: esso è inutilizzabile, in quanto il tempo di risposta è troppo elevato per avere un impiego pratico.

L'esposto seguente, tratto da un lavoro di Gary & Johnson, ne dà una fondamentale prova pratica, considerando un calcolatore le cui istruzioni durano 0,000001 s per essere processate; naturalmente l'algoritmo considera input variabili, di lunghezza 10, 20, 30, 40, 50 e 60 e si suppone, per semplicità, che la lunghezza dell'input determini in modo quanto più preciso e linearmente dipendente il numero delle operazioni che devono essere eseguite: pertanto, se l'input ha lunghezza 60 significa che sono necessarie 60 operazioni per terminare l'algoritmo e quindi 0,00006 s sarà il tempo impiegato per l'esecuzione.

Se, invece, la complessità dell'algoritmo è quadratica, allora ciò significa che se l'input è di lunghezza 60, il numero di operazioni diviene $60 \cdot 60 = 3600$ e quindi il numero di secondi diviene 0,0036 s.

Se la complessità è cubica, allora con lunghezza dell'input di 60 il numero di operazioni diviene $60 \cdot 60 \cdot 60$ e quindi il tempo impiegato è di 0,216 s.

Con una complessità quintica, a 60 di lunghezza corrispondono 60^5 istruzioni e quindi 13 minuti di esecuzione.

Passando ad una complessità esponenziale, come 2^n , con lunghezza dell'input pari a 60 si hanno 2^{60} istruzioni e quindi un tempo esecutivo di 360 secoli. Passando appena a 3^n , con lunghezza dell'input pari a 60 si hanno 3^{60} istruzioni e quindi un tempo esecutivo di $1,3 \cdot 10^{13}$ secoli.

Osservazione: Si osservi che considerare un calcolatore 1000 volte più veloce può essere significativo se la complessità dell'algoritmo è polinomiale, ma è totalmente ininfluente se la complessità è esponenziale.

4.1 Algoritmo di Euclide

Si consideri il seguente algoritmo, noto come **algoritmo di euclide**, estremamente fulmineo:

Algorithm 6 Euclide

```
1: begin
2:    $a, b = m, n$ 
3:   while  $b \neq 0$  do  $a, b = b, a \text{ mod } b$ 
4:   mcd = a
5: end
```

In questo caso l'algoritmo considera come input **due numeri interi** m e n di cui ha senso determinare l'**m.c.d.**: pertanto essi devono essere necessariamente interi per ipotesi, in quanto nell'algoritmo non è previsto un controllo sintattico a monte. Generalmente si considerano $m > n$, ma ciò è ininfluente, in quanto il programma provvede ad effettuare un cambiamento del loro ordine in automatico.

Nell'algoritmo vi sono delle assegnazioni composte, in cui

$$a, b = m, n$$

ovvero ad a si assegna il valore m , mentre a b si assegna il valore n . Così come in seguito si ha

$$a, b = b, a \text{ mod } b$$

ovvero ad a si assegna il vecchio valore b , mentre a b si assegna il resto della divisione intera tra a e b . Alla fine del ciclo si ottiene che l'**m.c.d.** cercato è proprio a .

Si consideri, a tal proposito, il seguente esempio:

$$\begin{array}{|c|c|} \hline 12 & 9 \\ \hline a & b \\ \hline \end{array}$$

Dopo il primo passo si ottiene

$$\begin{array}{|c|c|} \hline 9 & 3 \\ \hline a & b \\ \hline \end{array}$$

ed infine

$$\begin{array}{|c|} \hline 3 \\ \hline a & b \\ \hline \end{array}$$

ecco che nella prima cella si ha proprio l'm.c.d. cercato. Ora, tuttavia, bisogna verificare se tale algoritmo sia effettivamente corretto e che quella considerata non sia solo una combinazione; inoltre, per determinare la complessità dell'algoritmo è necessario considerare, essenzialmente, il numero di **iterazioni libere** del ciclo while, in quanto la complessità dell'algoritmo dipende unicamente da tale fattore.

Osservazione: Si osservi, innanzitutto, che il ciclo while si chiude visto che in posizione b sarà presente un resto, ovvero una quantità intera che progressivamente diminuisce fino a diventare 0. L'ultimo passaggio, naturalmente, è il più semplice, in quanto l'algoritmo, prima dell'ultima iterazione, si ritroverà sempre nella situazione desiderata:

$$\boxed{\alpha \cdot \text{m.c.d.}} \quad \boxed{\text{m.c.d.}}$$

per cui nell'ultima iterazione, effettuando la divisione intera tra una quantità e un suo multiplo non si può che ottenere resto 0 e, quindi, nella cella a si ha proprio l'm.c.d. che si sta cercando, come mostrato di seguito:

$$\boxed{\text{m.c.d.}} \quad \boxed{0}$$

Per la dimostrazione della correttezza dell'algoritmo, anche nelle fasi intermedie, si deve dimostrare essenzialmente che, a qualunque fase del processo esecutivo

$$\text{M.C.D.}(a, b) = \text{M.C.D.}(b, a \bmod b)$$

ovvero che quello che si trova nelle celle $a - b$ nell'istante t e ciò che si trova in $a - b$ all'istante successivo hanno entrambi lo stesso massimo comune divisore: bisogna, di fatto, dimostrare quella che in matematica prende il nome di **invariante**; in questo caso l'invariante è proprio l'm.c.d. Infatti, le quantità che si trovano progressivamente nelle due celle hanno sempre lo stesso massimo comune divisore, a qualunque istante, fino ad arrivare alla configurazione finale in cui nella prima cella è presente un multiplo del valore della seconda cella.

Per dimostrare, quindi, l'invarianza seguente:

$$\text{M.C.D.}(a, b) = \text{M.C.D.}(b, a \bmod b)$$

si deve dimostrare che un diviore di (a, b) è anche divisore di $(b, a \bmod b)$ e viceversa, per cui hanno lo stesso massimo comune divisore. Si supponga, allora, che α divida sia a che b , per cui si ha che $a = \alpha a'$ e $b = \alpha b'$; inoltre, considerando $r = a \bmod b$ quale il resto della divisione tra a e b , si può scrivere, per il teorema del quoziente e resto

$$a = qb + r$$

Pertanto, sapendo che α divide sia a che b , bisogna verificare che α divida anche $r = a \bmod b$ per concludere la dimostrazione dell'invarianza. Tuttavia, il teorema del quoziente e resto permette di ricavare r come segue in cui, ovviamente, si ha che

$$r = a - qb = \alpha \cdot (a' - qb')$$

in cui è evidente come α sia, effettivamente, un divisore anche di $r = a \bmod b$. Procedendo, ora, al contrario, supponendo che $b = \alpha b'$ e $r = \alpha r'$, sempre per il teorema del quoziente e resto si può scrivere $a = qb + r = \alpha \cdot (qb' + r')$, per cui, ancora una volta, α divide anche a , come volevasi dimostrare.

Osservazione: Avendo dimostrato la correttezza dell'algoritmo di Euclide, bisogna ora procedere alla verifica della sua straordinaria rapidità. Il problema della complessità dell'algoritmo di Euclide, il quale era un alessandrino, venne risolto brillantemente da un matematico francese di nome **Lamé**, nel 1844, il quale, a tutti gli effetti, è da reputarsi il padre della **teoria della complessità**, quando **Reynaud**, nel 1811 lo aveva già preceduto, ma si parla sempre della prima metà dell'800.

Per comprendere il meccanismo alla base dello studio della complessità, si consideri un numero rappresentato in base 10, quale il seguente $n = (37855)_{10}$, il quale viene progressivamente ridotto

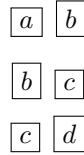
di **almeno** 10 volte ad ogni iterazione, portandolo progressivamente a 3785, 378, 37, 3 ed infine 0. Pertanto, al più, il numero di iterazioni necessarie per annientare questo valore e portarlo a 0 è pari alla **lunghezza decimale** del numero n (in questo caso pari a 5) che, con una buona dose di approssimazione può essere considerata

$$l_{10}(n) \cong \log_{10}(n)$$

Se, ora, il numero n considerato viene scritto in binario e ad ogni iterazione viene annientato della metà, al più serviranno un numero di iterazioni pari alla lunghezza binaria del numero n considerato per annientarlo, ovvero

$$\#\text{iterazioni} \leq l_2(n) \cong \log_2(n)$$

Con questa premessa, considerando l'algoritmo di Euclide e procedendo con i calcoli di **Lamé**, si prendano ad esempio tre iterazioni successive del ciclo while, che producono i seguenti tre stati corrispondenti a tre istanti consecutivi dell'esecuzione:



in cui, ovviamente, si ha che $a \geq b \geq c \geq d$; inoltre, per come sono stati ottenuti a, b, c e d appare evidente che

$$a = bq + c \text{ e } b = qc + d$$

A parte il caso iniziale in cui i due numeri $a \leq b$, cui l'algoritmo provvede cambiandoli d'ordine, in generale si ha sempre che $a > b$ e, quindi, il **quoziente** della divisione tra i due numeri è sempre **almeno 1**, quindi, siccome $q \geq 1$ deve essere che

$$b = qc + d \geq c + d$$

ma non solo: è anche noto che il valore della prima posizione è sempre maggiore o uguale del valore nella seconda posizione, ovvero $c \geq d$, per cui si ha che

$$b = qc + d \geq c + d \geq 2d$$

Questo significa che in ogni passo doppio, ovvero ogni due iterazioni, il contenuto della seconda cella di memoria è dimezzato, o peggio, in quanto si ha che $b \geq 2d$. Volendo far sì che l'elemento nella seconda cella divenga zero, poiché ad ogni doppia iterazione il suo valore viene almeno dimezzato, il numero di passi doppi necessari è al più uguale al logaritmo in base 2 di n , quindi

$$\#\text{passi doppi} \cong \log_2(n) \longrightarrow \#\text{passi singoli} < 2 \log_2(n)$$

pertanto, il numero di passi di cui si necessita per terminare il ciclo while è estremamente piccolo, anche nel caso in cui il numero n considerato è estremamente grande.

Osservazione: Un altro modo per descrivere l'algoritmo di Euclide è quello che riguarda la **ricorsività**, come mostrato di seguito

Algorithm 7 Euclide

```

1: Procedura Euclid(a,b)
2: if  $b = 0$  then
3:   return  $a$ 
4: else
5:   return Euclid( $b, a \bmod b$ )

```

Esempio: Per esempio, si considerino le iterazioni seguenti

$$\text{Euclid}(30, 21) \longrightarrow \text{Euclid}(21, 9) \longrightarrow \text{Euclid}(9, 3) \longrightarrow \text{Euclid}(3, 0) = 3$$

4.2 Notazione O grande

Si considerino due funzione $f(x)$ e $g(x)$ infinite a $+\infty$, ovvero

$$\lim_{x \rightarrow +\infty} f(x) = +\infty \quad \text{e} \quad \lim_{x \rightarrow +\infty} g(x) = +\infty$$

Allora le due funzioni si rassomiglieranno per $x \rightarrow +\infty$ quando hanno lo stesso ordine di infinito, ovvero

$$\lim_{x \rightarrow +\infty} \frac{f(x)}{g(x)} = \alpha \in \mathbb{R}^+ - \{0\}$$

Tuttavia, taluna è una generalizzazione molto spartana, ma la notazione O grande lo è ancora di più. Si considerino, nuovamente, due funzioni $f(x)$ e $g(x)$, le quali non debbono più essere infinite a $+\infty$, ma è sufficiente che esse siano **definitivamente positive**, ovvero

$$\exists x_n : \forall x > x_n, f(x) > 0 \wedge g(x) > 0$$

In questo caso, pertanto, affermare che $f(x)$ “assomiglia” a $g(x)$ significa affermare che $f(x)$ appartiene alla stessa **classe di equivalenza** di $g(x)$ (ovvero la classe delle funzioni che rassomigliano a $g(x)$), che si indica come segue

$$f(x) \in \Theta(g(x))$$

che, generalmente, verrà denotato con

$$f(x) = \Theta(g(x))$$

nonostante sia più propriamente corretto impiegare un segno di appartenenza in luogo di uno di uguaglianza.

APPARTENENZA ALLA CLASSE DI EQUIVALENZA DI UNA FUNZIONE

Si considerino due funzioni $f(x)$ e $g(x)$ **definitivamente positive**; si dirà che $f(x)$ appartiene alla stessa classe di equivalenza di $g(x)$ e si scriverà

$$f(x) = \Theta(g(x))$$

se

$$\exists c_1 > 0, c_2 > 0, \bar{x} : c_1 \cdot g(x) \leq f(x) \leq c_2 \cdot g(x), \quad \forall x \geq \bar{x}$$

ovvero si riesce a descrivere ipoteticamente, con la funzione $g(x)$, una guaina all'interno della quale racchiudere la funzione $f(x)$.

Osservazione: Dalla definizione appena fornita, appare evidente come questa sia a tutti gli effetti una **relazione di equivalenza**, in quanto è soddisfatta la proprietà **riflessiva**

$$g(x) = \Theta(g(x))$$

in quanto basta considerare $c_1 = c_2 = 1$; inoltre è soddisfatta anche la proprietà **simmetrica**:

$$f(x) = \Theta(g(x)) \longrightarrow g(x) = \Theta(f(x))$$

giacché se si ha che

$$c_1 \cdot g(x) \leq f(x) \leq c_2 \cdot g(x)$$

allora si può scrivere

$$\frac{1}{c_2} \cdot f(x) \leq g(x) \leq \frac{1}{c_1} \cdot f(x)$$

e infine quella **transitiva**

$$f(x) = \Theta(g(x)), g(x) = \Theta(h(x)) \longrightarrow f(x) = \Theta(h(x))$$

in quanto se

$$c_1 \cdot g(x) \leq f(x) \leq c_2 \cdot g(x) \quad \text{e} \quad c_3 \cdot h(x) \leq g(x) \leq c_4 \cdot h(x)$$

è facile concludere che

$$c_1 \cdot c_3 \cdot h(x) \leq c_1 \cdot g(x) \leq f(x) \quad \text{e} \quad f(x) \leq c_2 \cdot g(x) \leq c_2 \cdot c_4 \cdot h(x)$$

per cui si ottiene che

$$c_1 \cdot c_3 \cdot h(x) \leq f(x) \leq c_2 \cdot c_4 \cdot h(x)$$

Esempio: Si consideri la classe di equivalenza

$$\Theta(1)$$

Allora ad essa vi apparterranno tutte le costanti positive, così come le funzioni oscillanti che assumono valori positivi, come $f(x) = 2 + \sin(x)$, in quanto

$$2 \cdot 1 \leq f(x) \leq 3 \cdot 1, \quad \text{con } g(x) = 1$$

Osservazione: Si osservi che se due funzioni hanno lo stesso ordine di infinito, allora appartengono alla stessa classe di equivalenza. Tuttavia non è vero il contrario.

Si considerino, infatti, due funzioni $f(x)$ e $g(x)$ tali che

$$\lim_{x \rightarrow +\infty} \frac{f(x)}{g(x)} = \alpha \in \mathbb{R}^+ - \{0\}$$

quindi definitivamente, per $x \geq \bar{x}$, si ha che

$$\alpha - \epsilon \leq \frac{f(x)}{g(x)} \leq \alpha + \epsilon$$

dal momento che dalla definizione di limite sia $\forall \epsilon > 0$, è possibile anche considerare $\epsilon = \frac{\alpha}{2}$, da cui

$$\frac{\alpha}{2} \leq \frac{f(x)}{g(x)} \leq \frac{3}{2}\alpha$$

ma allora, moltiplicando ambo i membri per $g(x)$, essendo per definizione necessariamente definitivamente positiva, la diseguaglianza si conserva e diviene

$$\frac{\alpha}{2} \cdot g(x) \leq f(x) \leq \frac{3}{2} \cdot \alpha \cdot g(x)$$

che corrisponde esattamente alla definizione di due funzioni che appartengono alla stessa classe di equivalenza, con

$$c_1 = \frac{\alpha}{2} \quad \text{e} \quad c_2 = \frac{3}{2}\alpha$$

Per cui si è dimostrato che avere lo stesso ordine di infinito significa anche appartenere alla stessa classe di equivalenza, ma non è vero il contrario.

4.3 Funzioni di complessità

Si consideri una funzione di complessità temporale $f(n)$ definita in funzione n della lunghezza dell'input: è molto articolato, nonché scarsamente utile, provvedere al calcolo di valori precisi di una funzione di complessità temporale; presumibilmente, però, al crescere di n , ossia al crescere della lunghezza dell'input, aumenta anche il valore della funzione stessa.

Le funzioni $f(n)$ di complessità temporale oggetto di studio, comunque, sono molto generali e, comunemente, sono **definitivamente positive** e, generalmente, sono infinite per $x \rightarrow +\infty$ e vengono studiate per $x \geq 0$.

Com'è noto, è stata definita la classe di complessità $\Theta(f(n))$ come l'insieme di tutte le funzioni che rassomigliano alla funzione $f(n)$. In questo caso, in particolare, affermare che

$$g(n) = \Theta(f(n))$$

significa affermare che $\exists c_1 > 0, c_2 > 0, \bar{n}$ tali che

$$c_1 \cdot f(n) \leq g(n) \leq c_2 \cdot f(n), \quad \forall n \geq \bar{n}$$

in cui $\Theta(f(n))$ è, a tutti gli effetti, una **classe di equivalenza**, in quanto gode delle proprietà di **riflessività, simmetria e transitività**.

Se si considera una lunghezza $l = 5$ in base 10, una lunghezza effettiva senza zeri in testa, allora il numero n che si può descrivere con lunghezza $l = 5$ è compreso tra

$$10000 \leq n \leq 100000 - 1 \longrightarrow 10^{l-1} \leq n < 10^l$$

e passando ai logaritmi in base 10, ciò si traduce in

$$l_{10}(n) - 1 \leq \log_{10}(n) < l_{10}(n)$$

in cui $\log_{10}(n)$ viene considerata una **lunghezza continua**, decisamente più pratica nel suo utilizzo rispetto alla **lunghezza intera** $l_{10}(n)$, la quale è la lunghezza effettiva, ma inutilmente precisa. Naturalmente si può scrivere che

$$\frac{1}{2} \cdot l_{10}(n) \leq l_{10}(n) - 1 \leq \log_{10}(n) \leq l_{10}(n)$$

per cui si può osservare come $l_{10}(n)$ e $\log_{10}(n)$ appartengono alla stessa classe di equivalenza $\Theta(l_{10}(n))$ in quanto sono state usate le costanti $c_1 = \frac{1}{2}$ e $c_2 = 1$. Dal punto di vista della notazione Θ , quindi, tali lunghezze sono perfettamente equivalenti.

Non solo, ma se si considerano le quantità $\log_{10}(n)$ e $\log_2(n)$, è noto che

$$\log_2(n) = \frac{\log_{10}(n)}{\log_{10}(2)}$$

in cui

$$\frac{1}{\log_{10}(2)}$$

è una costante certamente positiva, allora si evince come tali quantità siano perfettamente equivalenti secondo la notazione Θ .

Quando si impiega la notazione Θ , infatti, una funzione $f(n)$ e $\alpha \cdot f(n), \alpha \geq 0$ appartengono alla stessa classe Θ ed è per questo che normalmente viene omessa la base dei logaritmi, che comunque deve essere maggiore di 1, la quale, di fatto, è ininfluente.

Inoltre si è osservato come due funzioni che hanno lo stesso ordine di infinito appartengono anche alla stessa classe Θ , semplicemente applicando la definizione di limite. Pertanto, le 3 funzioni

$$3n^2 + 1 \quad \frac{n^2}{1000} \quad 8945n^2 - n$$

appartengono alla stessa classe $\Theta(n^2)$, in quanto presentano lo stesso ordine di infinito, ovvero il limite del loro rapporto è una costante positiva.

Si consideri una funzione $f(n)$ infinita per $x \rightarrow +\infty$ e la funzione $g(n) = f(n) \cdot [\sin(n) + 2]$ anch'essa infinita per $x \rightarrow +\infty$ che appartengono alla stessa classe Θ ; tuttavia, si ha che

$$\lim_{n \rightarrow +\infty} \frac{g(n)}{f(n)} = \lim_{n \rightarrow +\infty} \sin(n) + 2 = \notin$$

per cui due funzioni che appartengono alla stessa classe di equivalenza Θ non è detto che abbiano lo stesso ordine di infinito.

FUNZIONE O-GRANDE DI UN'ALTRA

Si considerino due funzioni $f(x)$ e $g(x)$ **definitivamente positive**; allora indicare

$$f(n) = O(g(n))$$

significa affermare che definitivamente si ha che

$$f(n) \leq g(n)$$

ovvero che

$$\exists \bar{n}, c \geq 0 : f(n) \leq c \cdot g(n), \quad \forall n \geq \bar{n}$$

per cui se $f(n)$ e $g(n)$ appartengono alla stessa classe di equivalenza Θ , allora $f(n)$ è O-grande di $g(n)$ e $g(n)$ è O-grande $f(n)$ e viceversa: se $f(n)$ è O-grande di $g(n)$ e $g(n)$ è O-grande $f(n)$ allora $f(n)$ e $g(n)$ appartengono alla stessa classe di equivalenza Θ .

Esempio: Si osservi che, ovviamente

$$2n^2 = O(n^2) \quad \text{e} \quad \frac{n}{17} = O(n^2)$$

o ancora, che quando

$$\lim_{n \rightarrow +\infty} \frac{f(n)}{g(n)} = +\infty$$

significa affermare che $g(n) = O(f(n))$.

Osservazione: Si osservi, per quanto si è detto, che

$$\begin{cases} f(n) = O(g(n)) \\ g(n) = O(h(n)) \end{cases}$$

se e solo se appartengono alla stessa classe Θ . Analogamente si ha che

$$\begin{cases} f(n) = O(g(n)) \\ g(n) = O(h(n)) \end{cases}$$

allora si ha che $f(n) = O(h(n))$.

FUNZIONE Ω -GRANDE DI UN'ALTRA

Si considerino due funzioni $f(x)$ e $g(x)$ **definitivamente positive**; allora indicare

$$f(n) = \Omega(g(n))$$

significa affermare che definitivamente si ha che

$$f(n) \geq g(n)$$

ovvero che

$$\exists \bar{n}, c \geq 0 : f(n) \geq c \cdot g(n), \quad \forall n \geq \bar{n}$$

per cui se $f(n)$ e $g(n)$ appartengono alla stessa classe di equivalenza Θ , allora $f(n)$ è Ω -grande di $g(n)$ e $g(n)$ è Ω -grande $f(n)$ e viceversa: se $f(n)$ è Ω -grande di $g(n)$ e $g(n)$ è Ω -grande $f(n)$ allora $f(n)$ e $g(n)$ appartengono alla stessa classe di equivalenza Θ .

Esempio: Si osservi che indicare

$$f(n) = \Omega(g(n))$$

se e solo se

$$g(n) = O(f(n))$$

Analogamente si ha che

$$\begin{cases} f(n) = O(g(n)) \\ f(n) = \Omega(g(n)) \end{cases}$$

se e solo se appartengono alla stessa classe Θ .

4.4 Classi di complessità

Le tre classi Θ che si incontreranno maggiormente saranno

1. La classe **lineare** $\Theta(n)$;
2. La classe **log-lineare** $\Theta(n \cdot \log(n))$;
3. La classe **quadratica** $\Theta(n^2)$;

in cui si ha che

$$n = O(n \cdot \log(n)) \quad \text{e} \quad n \cdot \log(n) = O(n^2)$$

ovvero la complessità log-lineare è comunque poco più complessa rispetto a quella lineare, in quanto la crescita all'infinito è molto lenta.

Osservazione: Si osservi che la funzione

$$10^{10^{80}} n = \Theta(n)$$

anche se il suo coefficiente è estremamente elevato. Analogamente, se si ha una funzione

$$2^{0,00000\dots 1} = \Theta(2^n)$$

nonostante la sua crescita è estremamente lenta rispetto ad una funzione esponenziale normale. Tuttavia, tale risultato non costituisce un'anomalia nella notazione Θ , in quanto tali tipologie di funzioni sono impossibili tra trovare nella vita reale, secondo una valenza empirica.

Si consideri il listato seguente:

Algorithm 8 Esempio listato 1

```
1: for  $i = 1$  to  $k$ 
2:   for  $j = 1$  to  $k$ 
3:      $a_{i,j} = 0$ 
4:   next  $i$ 
5: next  $j$ 
```

Naturalmente il numero di iterazioni di tale algoritmo è $k \cdot k = k^2$, così come il numero delle assegnazioni. Ma se il listato fosse stato quello proposto di seguito:

Algorithm 9 Esempio listato 2

```
1:  $s = 0$ 
2: for  $i = 1$  to  $k$ 
3:   for  $j = 1$  to  $k$ 
4:      $s = s + a_{i,j}$ 
5:      $a_{i,j} = 0$ 
6:   next  $i$ 
7: next  $j$ 
8:  $s = s * s$ 
```

allora il numero della assegnazioni è significativamente aumentato, passando da k^2 a $2 + 2k^2$. Tuttavia, la notazione Θ fa sì che in ogni caso la complessità sia rimasta pari a k^2 , in quanto il numero di iterazioni, e quindi l'ordine, è **inalterato**.

Osservazione: Si osservi che nel caso dell'algoritmo **insertion-sort**, la complessità nel caso migliore è pari a $\Theta(n)$, in quanto il numero delle iterazioni è pari a $n - 1$, mentre nel caso peggiore la complessità è $\Theta(n^2)$, in quanto le iterazioni sono $\frac{n \cdot (n-1)}{2}$. Per quanto concerne la complessità tipica, ovverosia la complessità media, dal punto di vista empirico si osserva che essa è dell'ordine $\Theta(n^2)$.

Per quanto concerne l'algoritmo di Euclide, il numero di iterazioni, nel caso peggiore era di

$$1 + 2 \cdot \log_2(\min(m, n))$$

per cui si ha che la complessità effettiva sia $O(\log(n))$, ovvero possibilmente inferiore ad una complessità logaritmica.

4.5 Merge-Sort

L'algoritmo **merge-sort** è un algoritmo la cui idea alla base è molto più complessa dell'algoritmo **insertion-sort**.

L'input, come di consueto, è dato da un **array** A di lunghezza $\text{length}(A) = n$, per cui si dovranno ordinare n numeri. Inoltre, due altri input sono p e r , i quali costituiscono delle posizioni dell'array (anche intermedie o coincidenti), tali che $1 \leq p \leq r \leq n$.

Di seguito si espone il listato:

Algorithm 10 Merge-sort

```
1: MERGE-SORT( $A, p, r$ )
2: if  $p < r$ 
3:   then  $q = \left\lfloor \frac{p+r}{2} \right\rfloor$ 
4:   MERGE-SORT( $A, p, q$ )
5:   MERGE-SORT( $A, q+1, r$ )
6:   MERGE( $A, p, q, r$ )
```

Dove

$$\left\lfloor \frac{p+r}{2} \right\rfloor$$

prende il nome di *parte intera inferiore* della semisomma di p e r , ovvero viene eliminata la parte dopo la virgola.

Inizialmente l'algoritmo opera partendo ponendo $p = 1$ e $r = n$, ma successivamente le chiamate ricorsive fanno sì che p e r divengano delle posizioni intermedie, designate con q .

La procedura MERGE non è ricorsiva, ma presenta due cicli for in grado di ordinare due blocchi di numeri già ordinati. Tale procedura, infatti, inizia ad operare sui due insiemi di valori che sono già stati ordinati separatamente, con lo scopo di ricavare un unico insieme di valori ordinati, sfruttando l'ordinamento dell'insieme dei valori su cui si sta operando; per farlo si confrontano tutti i valori in testa agli insiemi ordinati, prendendo sempre il più piccolo tra i due che vengono confrontati, fino ad arrivare all'ultimo valore (necessariamente molto elevato, appositamente inserito come “fine-corsa”) che viene appositamente posto alla fine in modo da segnalare il termine dell'insieme di valori; in particolare si ha che $p < q < r$, in cui i valori che stanno tra $[p - q]$ e $[q - r]$ sono già stati ordinati.

La procedura viene esposta di seguito:

Algorithm 11 Merge

```

1: MERGE( $A, p, q, r$ )
2:  $n_1 = q - p + 1$ 
3:  $n_2 = r - q$ 
4: Create two new arrays  $L$  and  $R$ 
5: for  $i = 1$  to  $n_1$ 
6:   do  $L[i] = A[p + i - 1]$ 
7:      $R[j] = A[q + j]$ 
8: ...
9:  $L[n_1 + 1] = +\infty, R[n_2 + 1] = +\infty$ 
10:  $i = 1, j = 1$ 
11: for  $k = p$  to  $r$ 
12:   do if  $L[i] \leq R[j]$ 
13:     then  $A[k] = L[i]$ 
14:        $i = i + 1$ 
15:     else  $A[k] = R[j]$ 
16:        $j = j + 1$ 
```

In cui le iterazioni coinvolte, naturalmente, fino alla riga 10 sono

$$n_1 + n_2 = q - p + 1 + r - q = r - p + 1$$

ovverosia il numero delle posizioni che vanno da p a r .

Mentre dalla riga 10 in poi il numero di iterazioni è, ovviamente, ancora una volta $r - p + 1$; pertanto il numero di iterazioni complessive è pari al doppio di $r - p + 1$, quindi la complessità è dell'ordine

$$\Theta(r - p + 1)$$

Esempio: Si consideri il seguente array

$$\boxed{4} \ \boxed{5} \ \boxed{3} \ \boxed{1} \ \boxed{7} \ \boxed{2} \ \boxed{0} \ \boxed{9}$$

in cui $n = 8$. Per l'ordinamento di tale array, si divide lo stesso in due parti e poi successivamente ogni parte in ancora due parti e nuovamente in due parti fino ad ottenere una divisione dell'ordinamento riducendosi all'unità, in cui i singoli valori saranno automaticamente ordinati, come mostrato di seguito:

$$\boxed{4} \ \boxed{5} \ \boxed{3} \ \boxed{1} \mid \boxed{7} \ \boxed{2} \ \boxed{0} \ \boxed{9}$$

e poi

$$\boxed{4} \ \boxed{5} \mid \boxed{3} \ \boxed{1} \mid \boxed{7} \ \boxed{2} \mid \boxed{0} \ \boxed{9}$$

ed infine

$$\boxed{4} \mid \boxed{5} \mid \boxed{3} \mid \boxed{1} \mid \boxed{7} \mid \boxed{2} \mid \boxed{0} \mid \boxed{9}$$

Dal momento che i blocchi da 1 sono già ordinati si procede a ad utilizzare la procedura *Merge* per ordinare i blocchi da 2, ottenenendo

$$\boxed{4} \ \boxed{5} \mid \boxed{1} \ \boxed{3} \mid \boxed{2} \ \boxed{7} \mid \boxed{0} \ \boxed{9}$$

e poi procedendo con i blocchi da 4, applicando la procedura *Merge* si ottiene

$$\boxed{1} \ \boxed{3} \ \boxed{4} \ \boxed{5} \mid \boxed{0} \ \boxed{2} \ \boxed{7} \ \boxed{9}$$

e infine si ordina il blocco completo, applicando la procedura *Merge* si ottiene

$$\boxed{0} \ \boxed{1} \ \boxed{2} \ \boxed{3} \ \boxed{4} \ \boxed{5} \ \boxed{7} \ \boxed{9}$$

Questa, naturalmente, è una **esecuzione parallela**, mentre quella dell'algoritmo è **lineare**, in quanto va ad analizzare progressivamente ogni ramo di un albero binario, risalendolo progresivamente.

Osservazione: Tale algoritmo è **profondamente rigido**, in quanto il numero delle iterazioni che si devono eseguire è sempre lo stesso, indipendente dal caso migliore, peggiore e generale.

Questa non è una buona caratteristica, la quale rende ragionevole l'utilizzo di tale algoritmo quando i numeri sono inizialmente fortemente disordinati; se i numeri iniziali sono parzialmente ordinati è più conveniente impiegare l'algoritmo di insertion-sort.

La complessità di tale algoritmo, tuttavia, nel caso peggiore è **log-lineare** ed è estremamente ottima in quanto un algoritmo di ordinamento generale, come il merg-sort (che non richiede una specifica struttura dei dati in ingresso) non è in grado di battere una complessità log-lineare.

Il principio di costruzione di tale algoritmo, nonché la tecnica impiegata per la risoluzione di tale problema, prende il nome di *Divide et impera* (o *Divide and conquer*).

16 Marzo 2022

Una componente fondamentale dello studio algoritmico è la **relazione di ricorrenza lineare**, alla base degli algoritmi che si fondano sulla **ricorsività**.

È noto che la sezione aurea si basa sulla suddivisione di un segmento in maniera armoniosa:

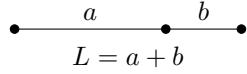


Figura 6: Sezione aurea

Rispettando la proporzione seguente

$$a + b : a = a : b$$

e ponendo $b = 1$ si ottiene che

$$a + 1 : a = a : 1$$

che dà vita all'equazione di secondo grado seguente (e ricordando che il prodotto dei medi è uguale al prodotto degli estremi):

$$a^2 - a - 1 = 0$$

che produce due soluzioni

$$a_{1,2} = \frac{1 \pm \sqrt{5}}{2}$$

in cui la soluzione più importante è quella naturalmente positiva, ovvero:

$$\phi = \frac{1 + \sqrt{5}}{2} \cong 1.62 \quad \text{e} \quad \psi = \frac{1 - \sqrt{5}}{2}$$

Esempio: Alla base dello studio della ricorsività vi è il *Liber Abaci* di Fibonacci, il quale è anche l'ideatore della sequenza di Fibonacci e delle straordinarie proprietà ad essa collegate; all'interno del *Liber Abaci* si espone anche il problema della crescita demografica dei conigli, alla base del quale vi è proprio una relazione di ricorrenza.

Per la visualizzazione del problema si parte dalla condizione iniziale che riguarda la popolazione dei conigli seguente:

$$F_1 = F_2 = 1$$

e si espone anche una legge che permette di calcolare l'evoluzione demografica dei conigli nel tempo, ossia la cosiddetta legge di Fibonacci, ovvero una **relazione di ricorrenza lineare**:

$$F_n = F_{n-2} + F_{n-1}$$

Naturalmente, nella risoluzione del problema, ciò che era maggiormente importante era determinare una formula chiusa, la quale è la seguente:

$$F_n = \frac{\phi^n - \psi^n}{\sqrt{5}}$$

ma siccome $\psi < 0$, implode a 0, per cui con ottima approssimazione si può affermare che

$$F_n = \frac{\phi^n - \psi^n}{\sqrt{5}} \cong \frac{\phi^n}{\sqrt{5}}$$

che sottolinea il carattere esponenziale della crescita dei conigli; tale risultato, inoltre, permette di ottenere anche l'evidenza seguente:

$$\lim_{n \rightarrow +\infty} \frac{F_{n+1}}{F_n} = \phi$$

Pertanto la relazione di ricorrenza lineare è accompagnata da una formula chiusa che, in questo caso, è una formula esatta (un risultato impressionante visto che, partendo da una relazione di ricorrenza, ottenere una formula chiusa non è banale).

In generale, infatti, un approccio sistematico ad una relazione di ricorrenza lineare permette di ottenere una formula chiusa, la quale, però, è tutt'altro che precisa, in quanto fornita in termini della notazione O-grande.

4.6 Master Theorem

Il **teorema maestro**, o il **teorema principale delle relazioni di ricorrenza**, si può applicare solamente in casi ben determinati (e particolarmente fortunati). In ogni caso, tale teorema inizia ad operare partendo dall'**espressione della condizione iniziale**

$$T(1) = \Theta(1)$$

ovvero $T(1)$ appartiene alla classe di equivalenza $\Theta(1)$, la quale è una definizione decisamente vaga (esistono due costanti positive tra cui $T(1)$ è compreso, ma null'altro), per cui non vale la pena di scriverla, giacché non interviene in nessuna maniera nella soluzione (benché sia fondamentale che vi sia, per decretare l'applicazione del teorema o meno).

La relazione, invece, alla base del teorema è la seguente

$$T(n) = a_1 \cdot T\left(\frac{n}{\lfloor b \rfloor}\right) + a_2 \cdot T\left(\frac{n}{\lceil b \rceil}\right) + f(n)$$

in cui $f(n)$ è una funzione nota. Tuttavia, la parte intera inferiore o superiore è ininfluente, per cui, posto $a = a_1 + a_2$, si preferisce la formula abbreviata

$$T(n) = a \cdot T\left(\frac{n}{b}\right) + f(n)$$

imponendo che $a \geq 1$ e $b > 1$ (altrimenti non è possibile applicare il teorema considerato, cercando di ricorrere ad altri metodi).

Supponendo che le condizioni di cui sopra siano verificati, è possibile applicare il teorema: l'applicazione del teorema maestro prevede 3 differenti casi, a seconda della **natura della funzione $f(n)$** :

- la funzione $f(n)$ è molto piccola
- la funzione $f(n)$ è giusta
- la funzione $f(n)$ è molto grande

Naturalmente, per confrontare tale funzione $f(n)$ se ne definisca una seconda, chiamata **funzione di confronto**:

$$\phi(n) = n^{\log_b(a)}$$

che assicura che la funzione sia crescente grazie alla base $b > 1$ del logaritmo. Pertanto si ha che il *master theorem* può essere applicato in questi tre differenti casi:

- La funzione $f(n)$ è **molto** più piccola della funzione $\phi(n)$, tuttavia **non è sufficiente** scrivere

$$f(n) = O(\phi(n))$$

per cui di fatto, rimane una sorta di “buco” in cui il teorema non si può applicare, ossia quando la funzione $f(n)$ è troppo piccola per rientrare nella seconda casistica, ma troppo grande per entrare nel primo caso.

- La funzione $f(n)$ rassomiglia alla funzione $\phi(n)$, ovvero

$$f(n) = \Theta(\phi(n))$$

- La funzione $f(n)$ è **molto** più grande della funzione $\phi(n)$, che è una condizione che in algoritmica serve a poco, in quanto non basta affermare che

$$f(n) = \Omega(\phi(n))$$

per cui, ancora una volta, rimane un buco, determinato da funzioni che sono troppo grandi per rientrare nella seconda casistica, ma troppo piccole per entrare nel terzo caso, per cui il teorema non si può impiegare.

Tuttavia, alla terza e ultima casistica si deve aggiungere anche un’ulteriore condizione che la funzione $f(n)$ deve soddisfare affinché si possa applicare il teorema; il problema, però, che si viene a determinare, è che la condizione in questione non è stabile rispetto alla classe di equivalenza Θ di appartenenza della funzione $f(n)$: è possibile che due funzioni $f(n)$ e $g(n)$ appartenenti alla stessa classe di equivalenza possano l’una soddisfare la condizione, la l’altra no, permettendo nel primo caso di applicare il teorema, nel secondo di non applicarlo.

Dal momento che in algoritmica la funzione $f(n)$ non si conosce bene giacché è nota solamente la classe Θ a cui appartiene, non è sempre possibile sapere se si può applicare il teorema o meno nel terzo caso.

Pertanto in algoritmica i casi più importanti sono il primo e il secondo; tuttavia, il caso che viene trattato nel dettaglio, però, è solo il secondo. Nel secondo caso è noto che $a \geq 1$, $b > 1$ e le due funzioni $f(n)$ e la funzione di confronto $\phi(n)$ appartengono alla stessa classe; da notare che vi è **transitività**, per cui il fatto di non conoscere $f(n)$ diviene irrilevante quando è nota la classe di equivalenza Θ di appartenenza della funzione $f(n)$: se $f(n)$ è equivalente a $\phi(n)$, qualunque funzione equivalente a $f(n)$ è equivalente a $\phi(n)$ per transitività.

Tramite procedimenti matematici omessi, si giunge alla soluzione matematica seguente

$$T(n) = \Theta(f(n) \cdot \log(n))$$

in cui la base del logaritmo è ininfluente dal punto di vista della notazione Θ .

Osservazione: Nell’algoritmo **merge-sort**, la procedura **merge** permette di ordinare dei valori partendo da due insiemi di valori già ordinati, ottenendo un ultimo insieme perfettamente ordinato. La procedura *merge* non è una procedura ricorsiva, mentre l’algoritmo *merge-sort* si basa sul meccanismo della ricorsività, andando a scorrere un albero binario da sinistra verso destra, risalendolo progressivamente (con una esecuzione lineare e sequenziale, per nulla parallela).

L’algoritmo è assolutamente corretto, dal punto di vista esecutivo, è ciò è evidente, anche solo dal punto di vista empirico; tuttavia, ciò che non appare evidente è la complessità dell’algoritmo considerato.

Si ripropone di seguito, a tal proposito, il listato dell’algoritmo **merge-sort**:

Algorithm 12 Merge-sort

```

1: MERGE-SORT( $A, p, r$ )
2: if  $p < r$ 
3:   then  $q = \left\lfloor \frac{p+r}{2} \right\rfloor$ 
4:   MERGE-SORT( $A, p, q$ )
5:   MERGE-SORT( $A, q+1, r$ )
6:   MERGE( $A, p, q, r$ )

```

Lo sforzo per ordinare n numeri, naturalmente, è dato dalla seguente semplice relazione

$$T(n) = T\left(\frac{n}{2}\right) + T\left(\frac{n}{2}\right) + f(n)$$

ovvero la somma dello sforzo per ordinare ciascuna metà più l’operazione di *merge* delle due metà, denotata dalla funzione $f(n)$, che, in forma più compatta, può essere scritta come segue

$$T(n) = 2 \cdot T\left(\frac{n}{2}\right) + f(n)$$

ed applicando tale formula a qualunque passo si ottiene, per esempio nel caso di $(\frac{n}{2})$

$$T\left(\frac{n}{2}\right) = 2 \cdot T\left(\frac{n}{4}\right) + f\left(\frac{n}{4}\right)$$

mentre nel caso di $(\frac{n}{4})$ si ha

$$T\left(\frac{n}{4}\right) = 2 \cdot T\left(\frac{n}{8}\right) + f\left(\frac{n}{8}\right)$$

e via dicendo, fintantoché non si arriva a 1. Inoltre, per quanto concerne la procedura *merge*, partendo da un array iniziale A , si creavano ulteriori due array, uno di destra R , di dimensione

n_1 , e uno di sinistra L , di dimensione n_2 , e per farlo si necessitava di $n = n_1 + n_2$ iterazioni (ovviamente pari alla somma del numero di valori che compongono il primo e il secondo array di lavoro). Per l'ordinamento di tutti i valori, ancora una volta, si necessitava di $n = n_1 + n_2$ (sempre pari al numero totale di valori da confrontare) iterazioni, per un totale di $2n$ iterazioni: pertanto la complessità della procedura *merge* è dell'ordine $\Theta(n)$ (in quanto il coefficiente 2 è totalmente ininfluente per la notazione Θ).

Pertanto, il problema da risolvere tramite il *teorema maestro* è il seguente

$$T(n) = 2 \cdot T\left(\frac{n}{2}\right) + \Theta(n)$$

in cui, evidente, se confrontata con la formula generale del *teorema maestro*, si evince che $a = b = 2$; inoltre si osservi che impiegare $\Theta(n)$ in luogo di $f(n)$ è una scorrettezza, giustificata dal fatto che non è nota la funzione precisa appartenente alla classe $\Theta(n)$. Per l'applicazione del *teorema maestro* è necessario considerare una funzione di confronto, come quella precedentemente introdotta

$$\phi = n^{\log_b(a)}$$

ma essendo $a = b = 2$ si ottiene che

$$\phi = n^{\log_2(2)} = n^1 = n$$

Ecco che allora la funzione $f(n)$ che non è nota, ma appartiene alla classe $\Theta(n)$ rassomiglia alla funzione di confronto; pertanto si può applicare il secondo caso del *teorema maestro* e anche la soluzione ad esso associata, ovvero

$T(n) = \Theta(f(n) \cdot \log(n))$

che nel caso qui considerato, essendo $f(n) = \Theta(n)$ si traduce in

$$T(n) = \Theta(n \cdot \log(n))$$

ovvero la **complessità di merge-sort** è log-lineare, la quale è una complessità generale e **profondamente rigida**, in quanto le operazioni che vengono svolte sono sempre le stesse; pertanto tale complessità si mantiene costante sia nel caso migliore, sia nel caso peggiore, sia un quello medio. Una complessità log-lineare è leggermente peggiore rispetto ad una complessità lineare, ma è comunque una complessità ottima: infatti, esiste un teorema che afferma che, asintoticamente, un **algoritmo di ordinamento generale** (ovvero un algoritmo che si propone di ordinare dei dati indipendentemente da come essi vengono strutturati) è destinato ad avere una complessità nel caso peggiore **almeno log-lineare** e ad avere una complessità nel caso medio **almeno log-lineare**, che sono esattamente le prestazioni che merge-sort garantisce.

Particolarmente infruttuosa è la complessità di merge-sort nel caso migliore, la quale è sempre log-lineare, mentre nel caso dell'algoritmo insertion-sort è lineare. Questo permette di affermare che l'algoritmo merge-sort risulta essere particolarmente utile nel suo utilizzo quando i valori da ordinare non sono già ordinati, ma totalmente in disordine; nel caso fossero parzialmente ordinati, sarebbe più conveniente impiegare insertion-sort.

4.7 Heap-sort

La struttura di grafo più comunque è quella di **albero**; in un grafo semplice è sempre possibile costruire un percorso chiuso, in cui partendo da un primo vertice, utilizzando almeno un arco, solamente archi in un solo senso (ovvero un arco serve per andare solamente da a a b e non viceversa) e senza archi multipli (ovvero non è possibile che vi siano più archi che collegano a e b) si ritorna al vertice di partenza, come mostrato di seguito:

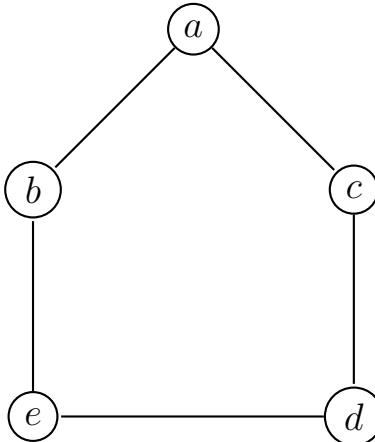


Figura 7: Esempio di grafo con percorso chiuso

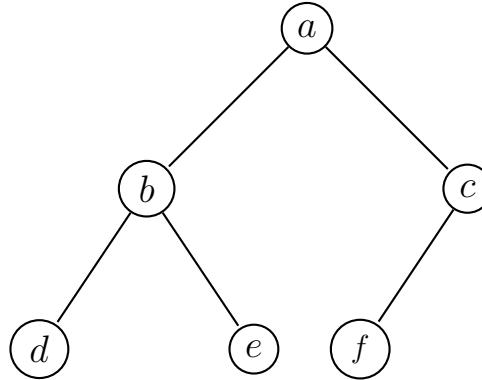
La definizione di albero segue dall'assenza di percorso chiusi all'interno di un grafo semplice, come esposto nella definizione che segue:

ALBERO

Un grafo semplice **privo percorsi chiusi** prende il nome di **albero**.

Tuttavia, tale definizione è eccessivamente generica, in quanto molto spesso si preferisce operare con **alberi radicati** (rooted-tree), ovvero degli alberi in cui è stata opportunamente specificata la **radice**, ossia un vertice specifico.

Un esempio di **albero radicato**, con **radice** (root) il vertice *a*, che verrà considerato viene raffigurato nel seguito:

Figura 8: Esempio di albero radicato di radice *a* e profondità 2

Dove per **profondità** è da intendersi il livello più basso raggiunto dai vertici dell'albero. I vertici che non presentano figli prendono, invece, il nome di **foglie** (in questo albero le foglie sono i vertici *d,e,f*).

L'albero radicato che maggiormente verrà considerato per la progettazione algoritmica è l'**albero binario**, in cui i figli di ogni vertice sono **al più 2**. Un albero binario ben strutturato prevede che vi siano *n* vertici e che l'albero sia **completo**, ovvero vengono saturati tutti i livelli di diversa profondità dell'albero prima di raggiungere al livello successivo, come mostrato di seguito:

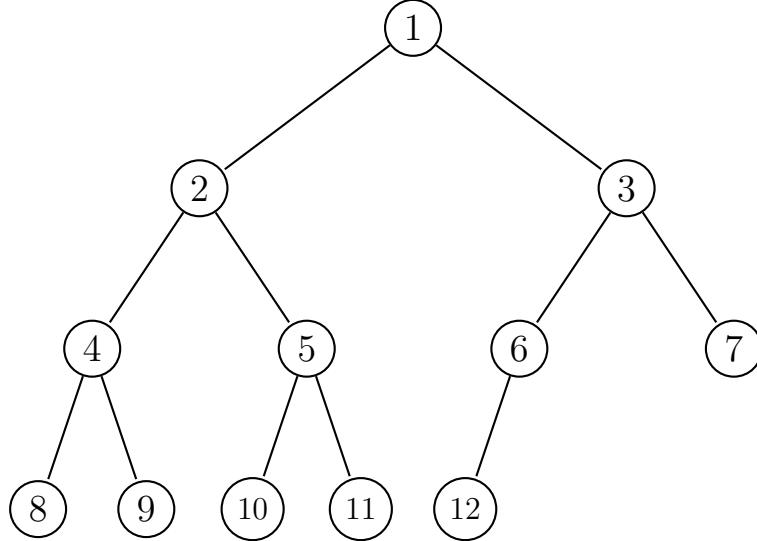


Figura 9: Esempio di albero binario completo

In cui la regola di completamento prevede di procedere da sinistra verso destra, completando ciascun livello prima di procedere in quello successivo, eccezion fatta per l'ultimo che, per ragioni di numero, non può essere completato. In un albero binario completo è facile capire la profondità dello stesso, semplicemente osservando che

$$2^h \leq n < 2^{h+1}$$

in cui h è la profondità dell'albero, mentre n è il numero di vertici dell'albero. Pertanto si ottiene che

$$h \leq \log_2(n) < h + 1$$

ovvero ciò significa che h e $\log_2(n)$ appartengono alla stessa classe Θ .

Dire che un albero è **super completo** significa che anche il livello delle foglie è pieno, e quindi n deve essere una potenza di 2 meno uno, ovvero $n = 2^h - 1$, con h l'altezza dello stesso. Inoltre è facile osservare come il numero di nodi in un albero **super completo** sia

$$2^0 + 2^1 + 2^2 + 2^3 + \dots + 2^{h-1} + 2^h = n$$

in cui h è la profondità dell'albero.

Osservazione: Si osservi la progressione geometrica seguente

$$1 + q + q^2 + q^3 + \dots + q^k = \sum_{i=0}^k q^i$$

alla quale può essere associata una formula chiusa, come mostrato di seguito

$$(1 + q + q^2 + q^3 + \dots + q^k) \cdot (q - 1) = (q + q^2 + \dots + q^{k+1} - 1 - q - q^2 - \dots - q^k) = (q^{k+1} - 1)$$

e dividendo ambo i membri per $(q - 1)$ si ottiene

$$(1 + q + q^2 + q^3 + \dots + q^k) \cdot (q - 1) = (q^{k+1} - 1) \longrightarrow \sum_{i=0}^k q^i = \frac{q^{k+1} - 1}{q - 1}$$

In cui, nel caso in cui la ragione $q = 2$, si ottiene che

$$\sum_{i=0}^k 2^i = 2^{k+1} - 1$$

Per cui è facile osservare che in un albero binario super completo, di altezza $h = k + 1$, la somma di tutti i vertici di ogni livello è pari a $2^h - 1$, come già dimostrato. Inoltre, nel livello delle foglie vi sarà sempre un numero di vertici, pari a 2^h , superiore a tutto il resto dell'albero, in cui il numero di vertici è $2^h - 1$, ovvero

$$\#\text{numero vertici} \cong 2 \cdot \#\text{numero foglie}$$

ovvero, dal punto di vista della notazione Θ , il numero di foglie è confondibile con il numero di vertici.

Osservazione: Il numero di archi in un albero di k vertici è esattamente pari a $k - 1$ (mentre nel caso di un grafo semplice, i nodi possono essere tutti isolati, con numero di archi pari a 0, oppure possono essere tutti connessi, con un numero di archi dell'ordine $\Theta(n^2)$), per cui dal punto di vista della notazione Θ il numero di vertici e il numero degli archi sono confondibili.

L'algoritmo **heap-sort** è un algoritmo di ordinamento per catasta, ove per catasta è da intendersi un albero binario completo, non necessariamente super completo, arricchito di ulteriori proprietà; tale proprietà all'interno della logica dell'**heap-sort**, è la designazione dei vertici, assegnando un valore numerico a ciascun vertice, facendo attenzione a far sì che il numero di ogni padre sia maggiore o uguale di ogni suo figlio (e di conseguenza, per ovvie ragioni di costruzione, di tutti quelli nei livelli sottostanti), ottenendo una **catasta** (heap), come mostrato di seguito:

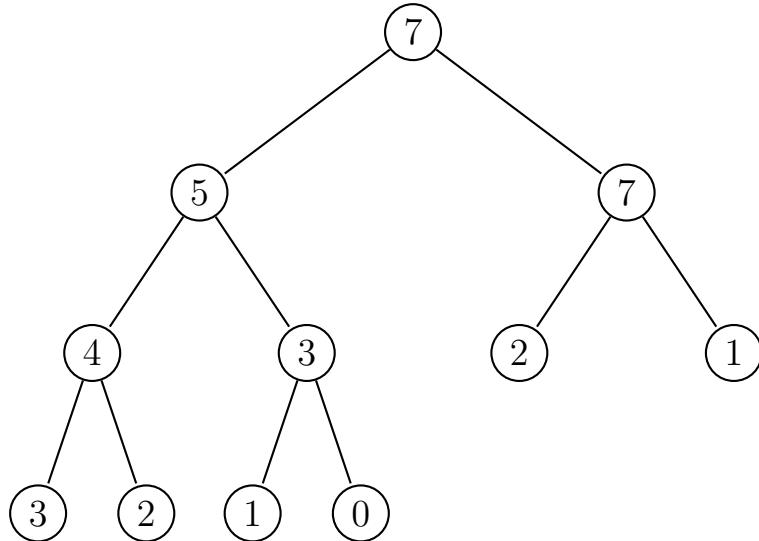


Figura 10: Esempio di catasta (heap)

All'interno della logica dell'**heap-sort** è possibile distinguere tra *maximum-heap* e *minimum-heap*: nel primo caso è il padre ad avere il valore numerico maggiore rispetto ai figli, nel secondo il genitore deve essere dominato da entrambi i figli.

Il primo algoritmo che si andrà a considerare prende il nome di **build-heap**, ovvero un algoritmo che è capace di creare una catasta che rispetta il vincolo del *maximum-heap* partendo da un albero in cui i nodi sono già pesati.

17 Marzo 2022

Si consideri un array, ovverosia una struttura dati assolutamente fondamentale in informatica, la quale presenta un numero di locazioni di memoria consecutive pari alla sua lunghezza $\text{length}(A) = n$.

Pertanto, scrivere

$A[3]$

significa considerare il contenuto della 3 cella di memoria consecutiva memorizzata nell'array A ; inoltre, si ha che lo sforzo per accedere ad una precisa locazione di memoria è dell'ordine $\Theta(1)$ (pari ad una sola operazione macchina), ovvero viene eseguito un **accesso diretto** (o **random access**, accesso casuale, o **accesso arbitrario**) alla cella di memoria richiesta, indipendentemente dalla lunghezza dell'array.

Un array è, quindi, una struttura a random-access, ovvero ad accesso lineare, designato con una lettera maiuscola (come A) e di una lunghezza n . Tuttavia, la sua struttura lineare si può interpretare come un albero binario (bidimensionale). Per esempio, l'array seguente

$$A = \begin{bmatrix} 7 & 8 & 2 & 9 & 4 & 1 & 0 & 5 & 6 \end{bmatrix}_{1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9}$$

viene rappresentato tramite il seguente albero binario

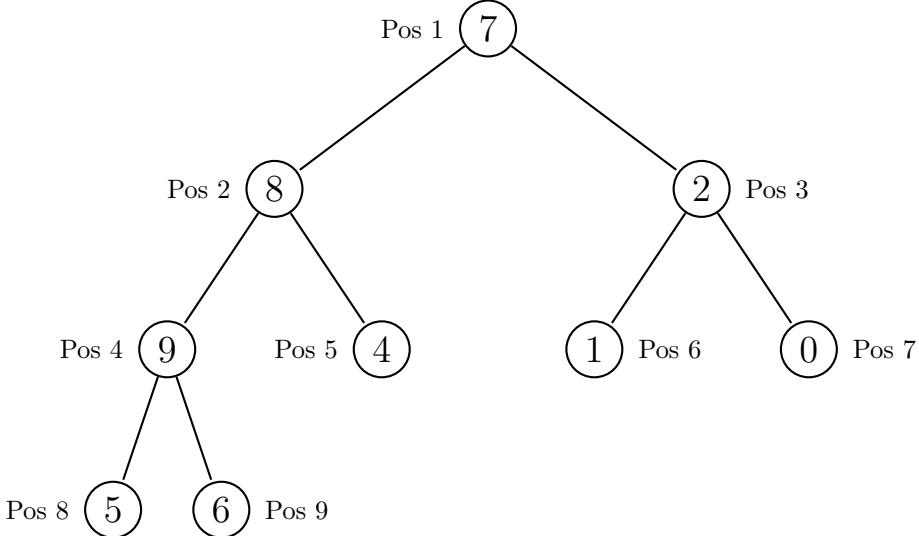


Figura 11: Esempio di catasta (heap)

La possibilità di considerare una struttura monodimensionale come un albero bidimensionale avviene tramite due algoritmi fondamentali (**heapify** e **buil-heap**), che adoperano le seguenti tre procedure:

Algorithm 13 Parent

1: PARENT(i)
 2: **return** $\lfloor \frac{i}{2} \rfloor$

Algorithm 14 Left

1: LEFT(i)
 2: **return** $2i$

Algorithm 15 Right

1: RIGHT(i)
 2: **return** $2i + 1$

In cui sono stati omessi degli opportuni controlli di validità dell'input i (come se $i = 1$ e si richiamasse la procedura Parent(1), la quale restituirebbe il valore 0 che è ovviamente corretto, o se non esiste un nodo sinistro/destro e si richiama la procedura Left/Right); questo in quanto si ha piena fiducia nell'utilizzo affidabile e consono di tali procedure; tramite esse, il calcolatore riesce ad interpretare una struttura monodimensionale come un albero bidimensionale.

La prima procedura che verrà impiegata nell'heap-sort (ordinamento per catasta) prende il nome di **heapify**, un sotto-algoritmo che verrà impiegato frequentemente dall'algoritmo **heap-sort**: una catasta, naturalmente, è una struttura ad albero in cui nei diversi nodi si trovano i numero che debbono essere ordinati (un ordinamento che poi si rifletterà nell'array di partenza).

Heapify opera su un sottoalbero dell'albero principale, la cui radice è uno dei nodi dell'albero primario individuato: ogni nodo, quindi, andrà ad individuare un sotto-albero (nel caso di una foglia, naturalmente il sottoalbero non esiste).

L'algoritmo **heapify**, tuttavia, viene utilizzato solamente quando il sottoalbero è noto a priori essere una catasta che rispetta il vincolo *maximum-heap*, eccezion fatta per la sottoradice del sottoalbero, in cui non è detto che venga rispettato tale vincolo.

Per comprendere il funzionamento di **heapify** si consideri il listato esposto di seguito, in cui si prende in considerazione il vertice di numero d'ordine i e il sottoalbero di A che presenta come sottoradice il nodo i :

Algorithm 16 Heapify

```
1: HEAPIFY( $A, i$ )
2:  $l = \text{LEFT}(i)$ ,  $r = \text{RIGHT}(i)$ 
3: if  $l \leq \text{heap-size}[A] \wedge A[l] > A[i]$ 
4:   then
5:     largest =  $l$ 
6:   else
7:     largest =  $i$ 
8: if  $r \leq \text{heap-size}[A] \wedge A[r] > A[\text{largest}]$ 
9:   then
10:    largest =  $r$ 
11: if largest  $\neq i$ 
12:   then
13:     Exchange:  $A[i] \leftrightarrow A[\text{largest}]$ 
14:     HEAPIFY( $A, \text{largest}$ )
```

in cui il termine “heap-size” fa riferimento alla dimensione, o altezza, del sottoalbero (ovvero al numero di nodi che costituiscono il sottoalbero).

La procedura esecutiva di heapify prevede dei controlli sequenziali: se esiste il figlio sinistro ($l \leq \text{heap-size}[A]$) e si verifica che il valore del figlio sinistro è maggiore del nodo radice ($A[l] > A[i]$), ponendo il valore maggiore all'interno di una posizione di lavoro denominata *largest*.

Dopodiché si deve procedere alla valutazione anche del figlio destro, se esiste ($r \leq \text{heap-size}[A]$), e si deve controllare se esso ha un valore maggiore sia del padre che del figlio sinistro ($A[r] > A[\text{largest}]$): se questo è il caso, allora all'interno di *largest* viene posta la posizione del figlio destro, ossia r .

Pertanto, ora, all'interno della posizione di lavoro *largest* vi è la posizione del padre, del figlio sinistro o del figlio destro, a seconda di chi abbia il valore massimo.

Naturalmente, ora, si procede con una modifica dell'albero solamente se all'interno di *largest* non vi è i , ossia la posizione del padre; se, infatti, vi fosse una posizione diversa da i significherebbe che uno dei due figli presenta un valore maggiore del padre e, quindi, il vincolo *maximum-heap* non sarebbe rispettato.

Quando si effettua uno scambio tra radice e figlio, ovviamente, è necessario ripetere la procedura **heapify** sul sottoalbero che presenta come sottoradice il figlio che aveva il valore maggiore, in quanto non è detto che taluno sia ancora una catasta.

Si consideri un sottoalbero come il seguente, in cui è possibile impiegare **heapify**, ovvero in cui la radice non è detto che domini i figli, ma il resto dell'albero rispetta il vincolo, come mostrato di seguito:

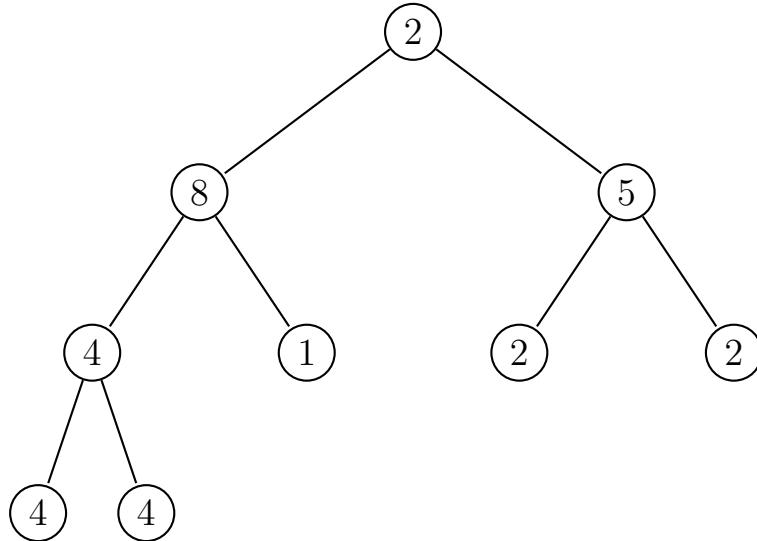


Figura 12: Esempio di sottoalbero al quale applicare heapify

Supponendo che il nodo radice di tale sottoalbero abbia il numero d'ordine $i = 3$, applicando l'algoritmo **heapify**, si considerano tre posizioni

6	7	$6 \text{ o } 3 \text{ o } 7$
l	r	largest

Naturalmente $l = 6$ in quanto la procedura $\text{LEFT}(i)$ restituisce $2i$, mentre $r = 7$ in quanto la procedura $\text{RIGHT}(i)$ restituisce $2i + 1$.

Operando secondo la procedura **heapify**, se il figlio sinistro esiste e il suo valore è maggiore di quello della sua radice, allora all'interno di *largest* andrà posto il valore $l = 6$, altrimenti il valore $i = 3$; Analogamente, se il valore del figlio di destra è maggiore di quello di sinistra e della radice, allora in *largest* andrà inserito il valore $r = 7$.

Nella posizione finale “*largest*”, pertanto, vi può essere la posizione del padre, del figlio sinistro o del figlio destro; se nella posizione “*largest*” vi è un valore più grande della radice, avviene l'**exchange**, come in questo caso, ottenendo

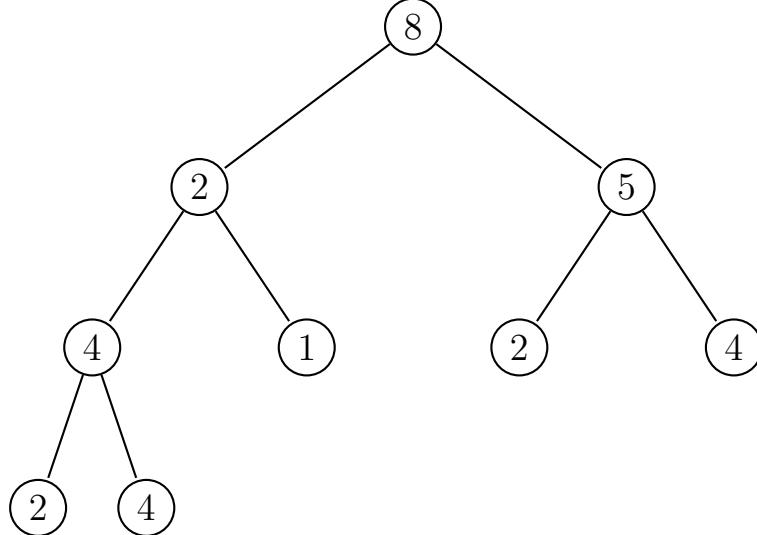


Figura 13: Esempio di sottoalbero al quale applicare heapify dopo un primo passo

Avendo operato tale sostituzione sul figlio di sinistra, è opportuno verificare il rispetto del vincolo del sottoalbero di sinistra (mentre il sottoalbero di destra non presenta anomalie, in quanto se

prima rispettava il vincolo, lo rispetta anche ora). Si applica, quindi, ancora una volta heapify considerando come sottoradice il nodo avente numero d'ordine 6:

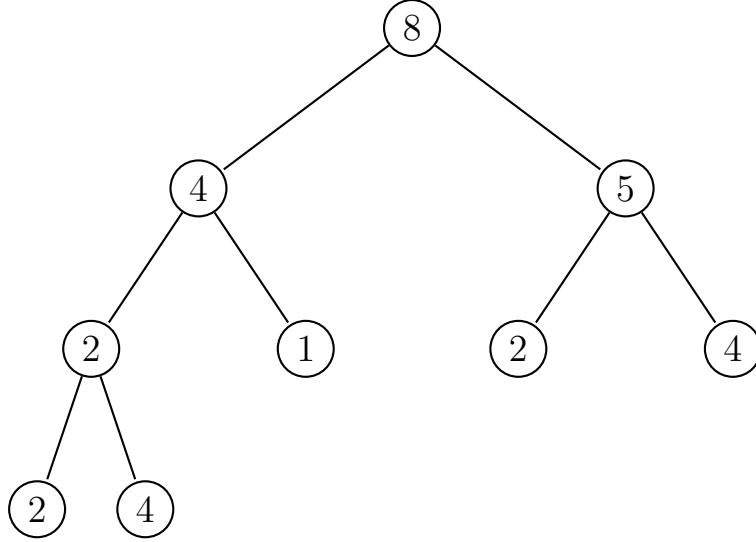


Figura 14: Esempio di sottoalbero al quale applicare heapify dopo un secondo passo

Ancora una volta, avendo scambiato la radice con uno dei figli, è necessario applicare heapify, questa volta considerando il sottoalbero avente sottoradice il nodo con numero d'ordine 12:

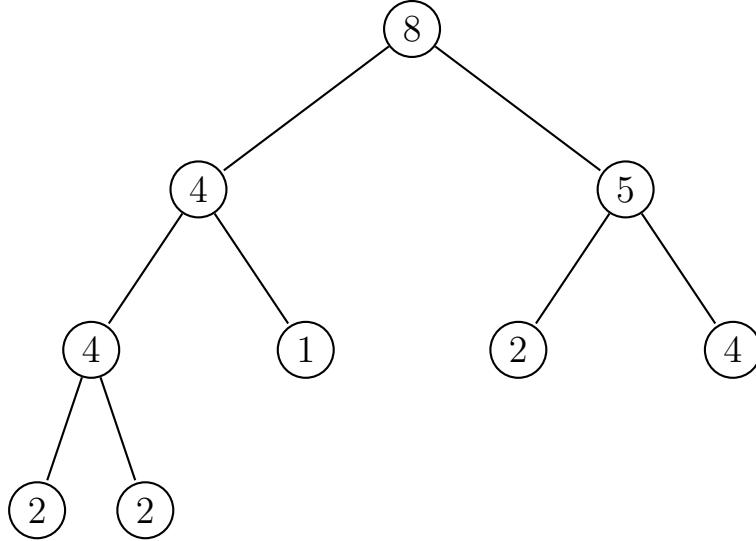


Figura 15: Esempio di sottoalbero al quale applicare heapify dopo un secondo passo

Avendo sostituito la radice con il figlio destro, è necessario applicare heapify ancora una volta, sul sottoalbero di destra; tuttavia, in questo caso, heapify non esegue alcuna operazione, giacché il sottoalbero di destra non esiste.

Ecco che l'albero iniziale, ora, rispetta il vincolo di *maximum heap*.

Osservazione: Si osservi che non sono presenti cicli in tale algoritmo, in quanto la complessità è data dalla ricorsività; in particolare, il numero delle iterazioni è

$$\#\text{iterazioni} \leq h_i + 1$$

in cui h_i è l'altezza del sottoalbero (infatti, nel caso peggiore, si deve scandagliare un albero per ogni livello, più una chiamata finale per constatare che non vi è un sottoalbero del nodo figlio);

pertanto la complessità è, operando una forte maggiorazione, dell'ordine $O(\log(n))$, in cui n è il numero dei nodi dell'albero intero (avendo interpretato il sottoalbero come l'albero intero); infatti è noto che la profondità di un albero completo avente n nodi è circa $\log(n)$, per cui si ha che h e $\log(n)$ appartengono alla stessa classe Θ .

Si osservi che, nel caso in cui si consideri l'albero intero di partenza, con H la sua altezza complessiva, allora si ha che

$$\#\text{iterazioni} \leq H + 1$$

Tuttavia, gli algoritmi più importanti da dover considerare sono altri due: **heap-sort** e **build-heap**, l'ultimo del quale serve proprio per costruire una catasta partendo da un albero non catastizzato, come mostrato di seguito:

Algorithm 17 Build-Heap

```
1: BUILD-HEAP( $A$ )
2: heap-size =  $\text{length}[A]$ 
3: for  $i = \lfloor \frac{n}{2} \rfloor$  down to 1
4:   do HEAPIFY( $A, i$ )
```

In cui

$$i = \left\lfloor \frac{n}{2} \right\rfloor$$

è il primo nodo con figli (in quanto è noto che il livello delle foglie, in un albero binario super completo, contiene più nodi di tutto il resto dell'albero) e che, quindi, potrebbe non rispettare il vingolo di *maximum-heap* (i nodi successivi, naturalmente, non possono avere di questo problema, visto che non presentano figli). Si consideri il caso seguente:

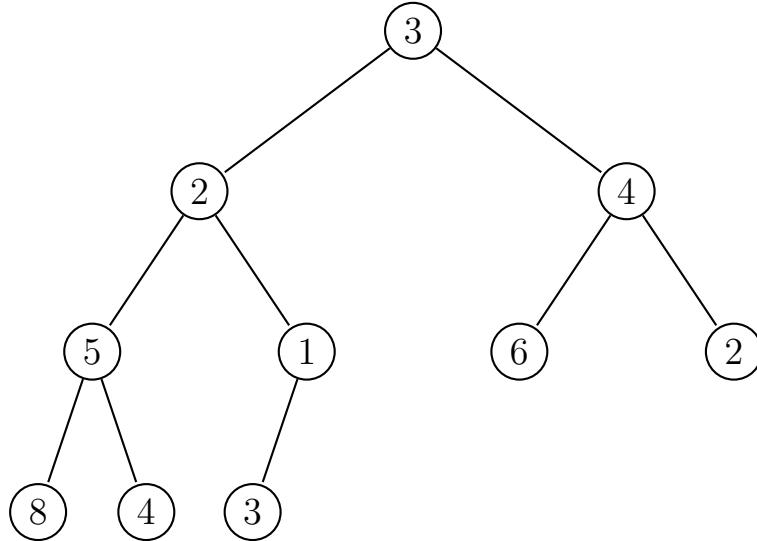


Figura 16: Esempio di albero da catastizzare

Quest'ultimo è un albero binario completo composto da 10 nodi. Pertanto il primo nodo che potrebbe avere figli è il noto 5, mentre le foglie non hanno alcun problema. A partire dal nodo avente numero d'ordine 5 si procede risalendo l'albero e applicando heapify, come mostrato di seguito, scambiando il nodo 5 con il nodo 10:

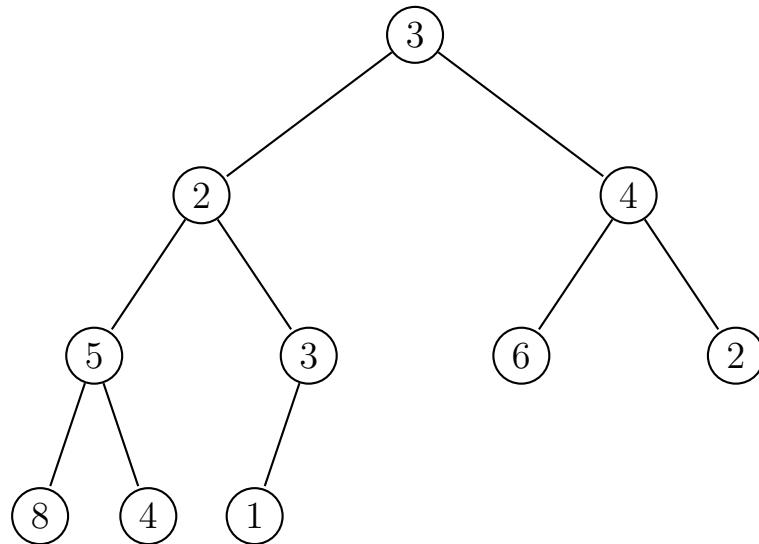


Figura 17: Esempio di albero da catastizzare dopo un primo passo

Risalendo progressivamente si considera il nodo precedente, ossia il 4 e si applica heapify, scambiando il nodo 4 con il nodo 8:

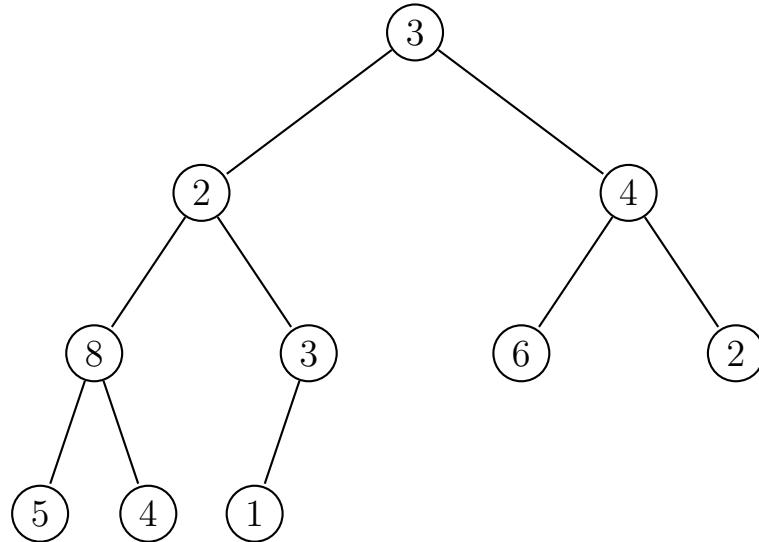


Figura 18: Esempio di albero da catastizzare dopo un secondo passo

Procedendo a ritroso si deve applicare nuovamente heapify, considerando il nodo precedente, ossia il 3 ed effettuando lo scambio tra il nodo 3 e il nodo 6:

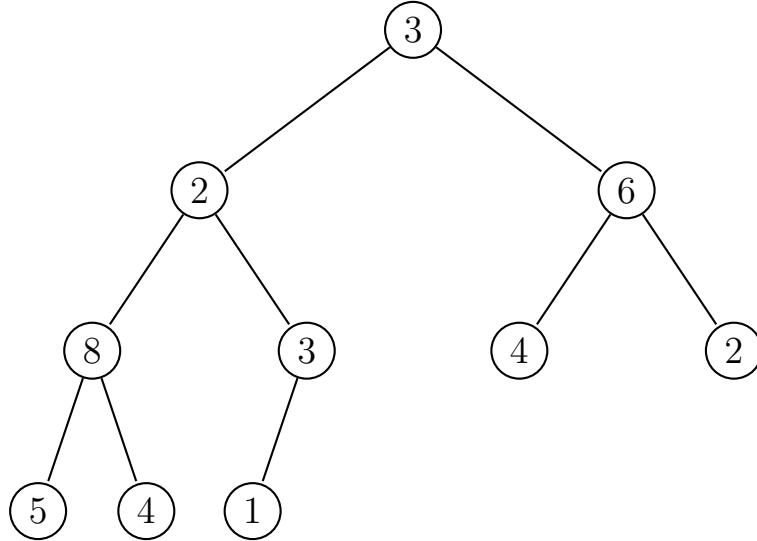


Figura 19: Esempio di albero da catastizzare dopo un terzo passo

Considerando ancora il nodo 2 si applica heapify e si scambia il nodo 2 con il nodo 4; tuttavia, heapify è ricorsivo e deve procedere operando sul sottoalbero avente sottoradice il nodo 4, effettuando uno scambio tra il nodo 4 e il nodo 8:

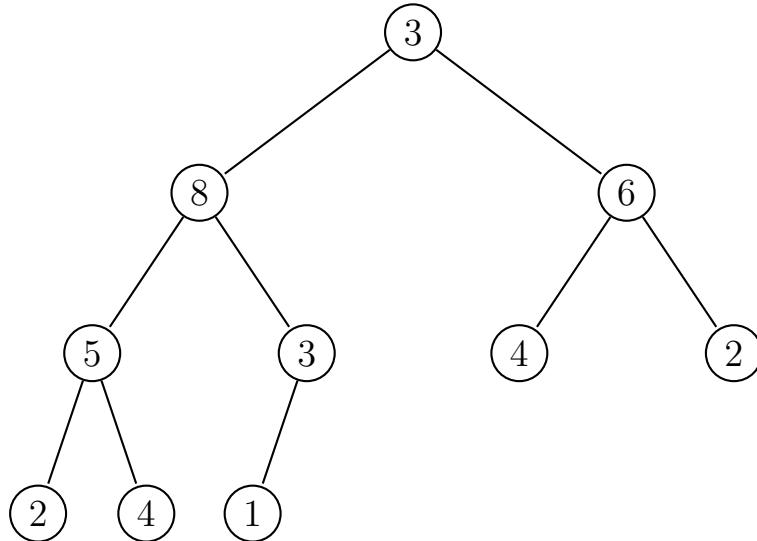


Figura 20: Esempio di albero da catastizzare dopo un quarto passo

Ed infine si applica heapify sul nodo 1, scambiandolo con il nodo 2; ancora una volta heapify deve scendere verso il basso, facendo rispettare il vincolo a tutti i sottoalberi sottostanti, scambiando il nodo 2 con il nodo 4 ed il nodo 4 con il nodo 9:

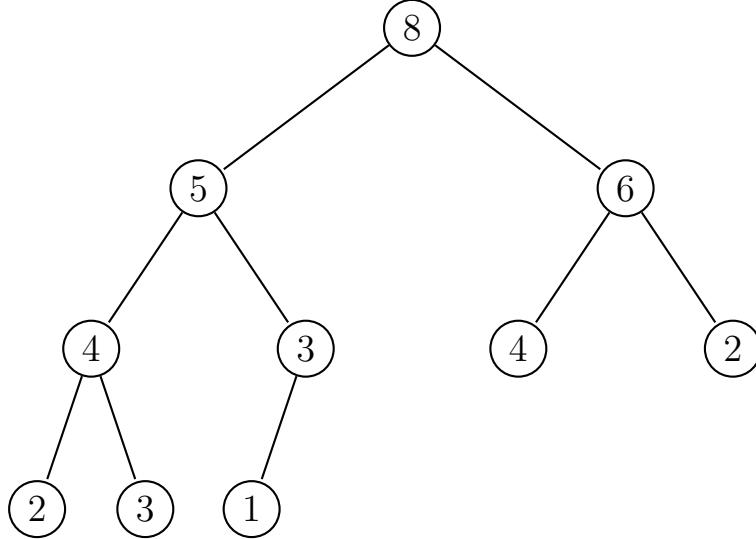


Figura 21: Esempio di albero da catastizzare dopo un quinto passo

Procedendo in questo modo, si procede alla creazione di una catastia, con un numero di iterazioni che dipende dalle chiamate ricorsive della procedura `heapify`; in particolare, si può osservare che il numero di nodi considerati da `build-sort` è approssimabile a

$$\frac{n}{2}$$

in quanto l'ultima metà è certamente esclusa, essendo i nodi delle foglie. Pertanto il ciclo for dell'algoritmo ha un numero di iterazioni pari a $\frac{n}{2}$; tuttavia, in ognuna di tali iterazioni viene richiamata la procedura `heapify`, la quale è ricorsiva e presenta una complessità di $O(\log(n))$; da ciò si evince che la complessità dell'algoritmo **build-heap** è pari a

$$O\left(\frac{n}{2} \cdot \log(n)\right) \cong O(n \cdot \log(n))$$

Inoltre è immediato capire che siccome il numero dei nodi che bisogna visitare in **build-heap** è pari a $\frac{n}{2}$, indipendentemente che `heapify` sia operativa o meno, si capisce che la complessità di **build-sort** è almeno lineare, ovvero

$$\Omega\left(\frac{n}{2}\right) \cong \Omega(n)$$

Quindi la complessità di **buil-heap** è almeno **lineare** e al più **log-lineare**; tuttavia, è sufficiente osservare che i nodi che la procedura `heapify` considera sono prossimi alle foglie (sempre perché considera la metà dei nodi in su ed è noto che se un albero binario è super completo, in livello delle foglie contiene più nodi di tutto il resto dell'albero), quindi i sottoalberi che vengono analizzati hanno tutti una lunghezza molto ridotta, che può abbassare la limitazione superiore assunta. Infatti, si può dimostrare che la complessità di calcolo di **build-heap** è lineare

$$\Theta(n)$$

e questo risultato, dal punto di vista dell'ordinamento, è totalmente irrilevante, anzi deve essere così affinché la complessità generale di altri algoritmi che ne fanno uso sia quantomeno accettabile.

Per conoscere come opera **heap-sort** si consideri l'albero binario seguente:

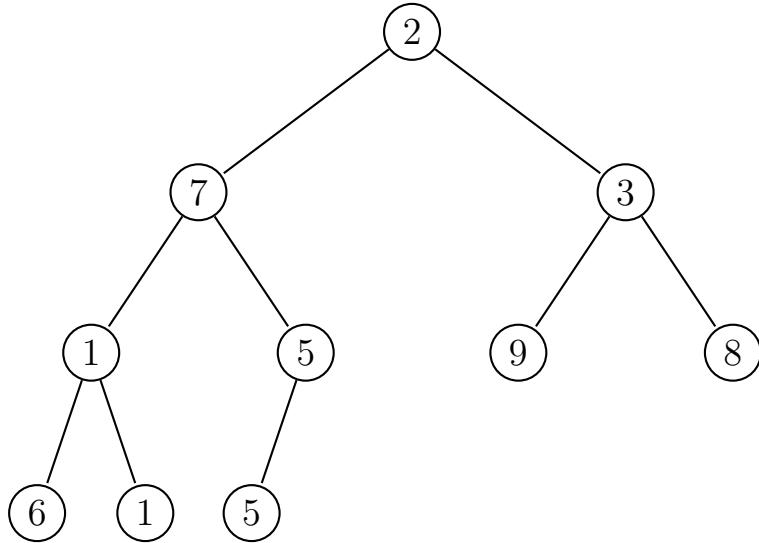


Figura 22: Esempio di albero da ordinare con heap-sort

Lavorando in loco, la prima operazione da eseguire su tale albero è quello di catastizzarlo, ottenendo una struttura completamente opposta rispetto a quella che si necessita: sul nodo radice è presente il valore maggiore, e sull'ultima posizione il valore più piccolo, com'è evidente nel seguente:

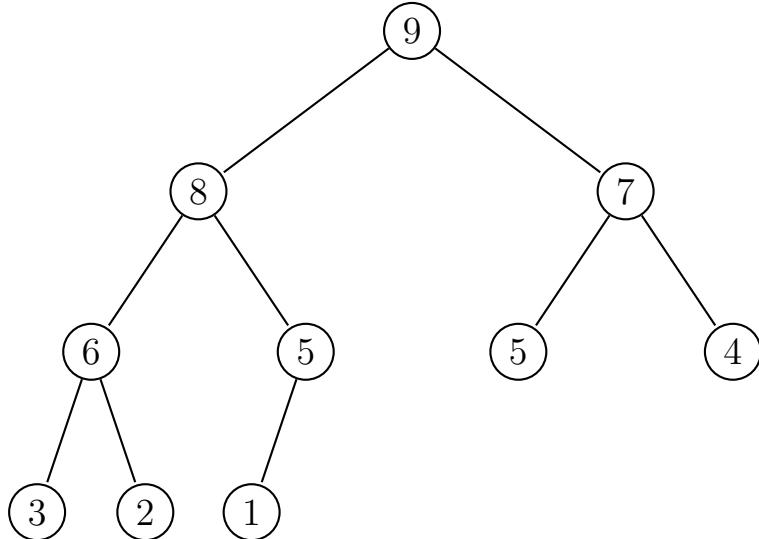


Figura 23: Esempio di albero catastizzato

Allora, secondo la logica heap-sort, si procede a sostituire il primo nodo con l'ultimo (effettuando un primo ordinamento); dal momento che l'ultimo nodo è un nodo foglia, non vi è problema che rispetti il vincolo *maximum-heap*, quindi si opera una cancellazione dell'ultimo nodo e dell'arco che lo congiunge con il nodo padre ($\text{heap-size}[A] = \text{heap-size}[A] - 1$).

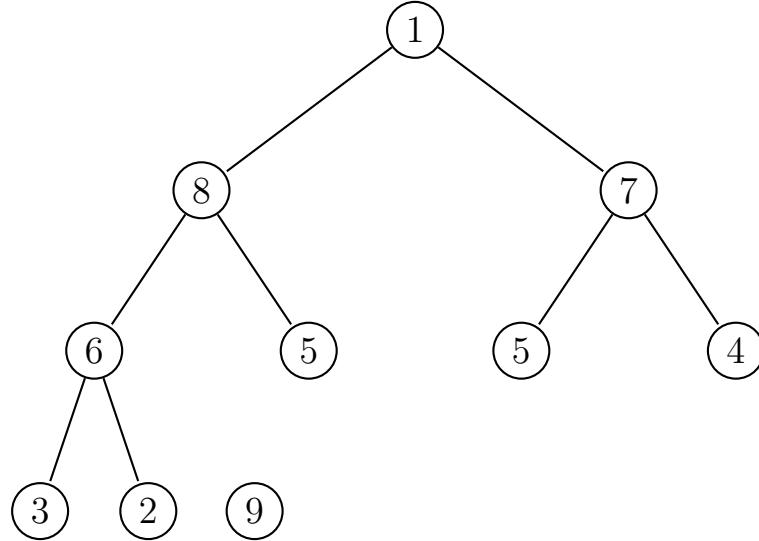


Figura 24: Esempio di albero da ordinare dopo un primo heap-sort

Ora si procede nuovamente a catastizzare, applicando heapify all'albero intero: da notare che è possibile applicare tale algoritmo in quanto l'albero intero è una catasta, essendolo già da prima; l'unico problema nel rispetto del vincolo, che giustifica l'applicazione di heapify, sta proprio nel nodo radice dell'albero intero:

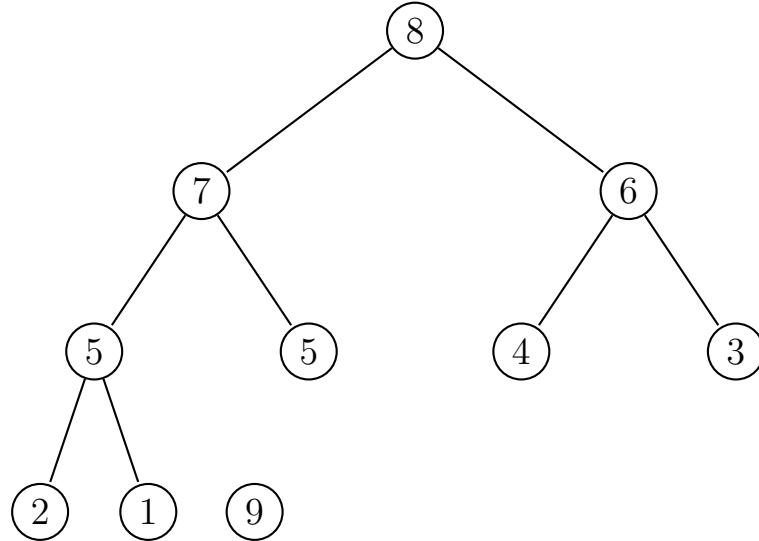


Figura 25: Esempio di albero catastizzato dopo un primo heap-sort

Ancora una volta si è ottenuta una catasta nella quale, dei valori rimasti, il maggiore si trova in prima posizione e il minore in ultima. Si procede, allora, a sostituire il primo nodo con l'ultimo, il quale viene eliminato insieme al suo collegamento con il nodo padre.

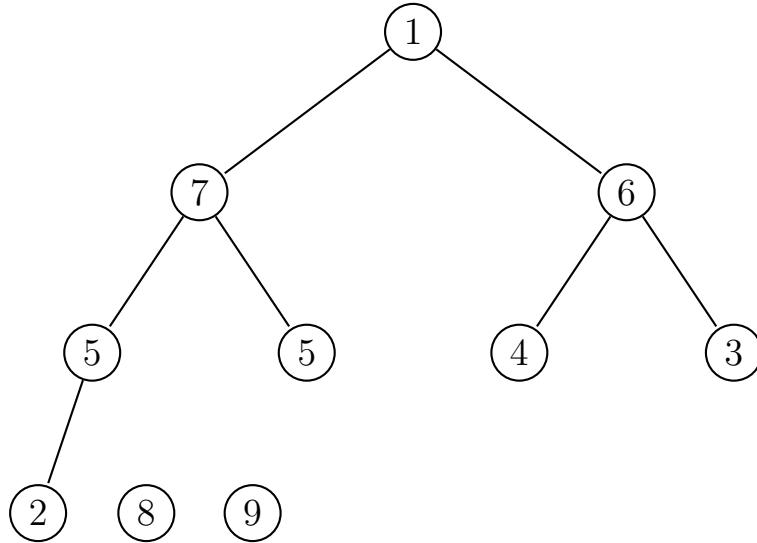
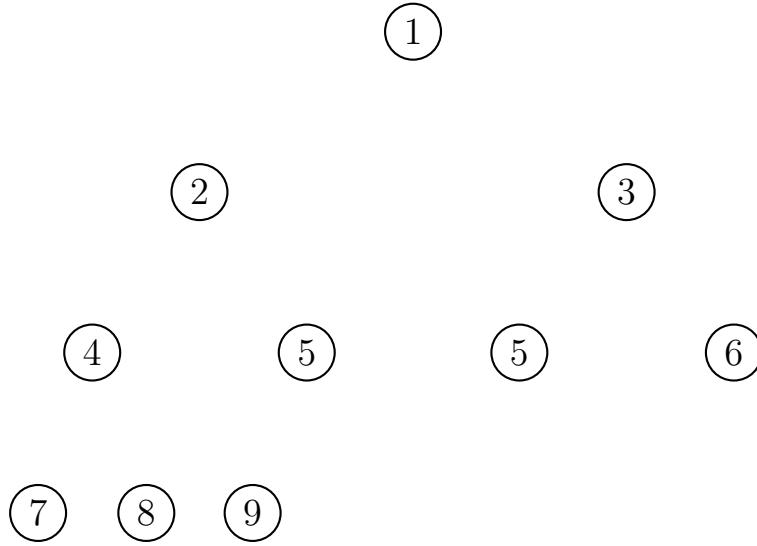


Figura 26: Esempio di albero da ordinare dopo un secondo heap-sort

Ecco che allora il penultimo nodo si trova perfettamente ordinato con gli altri; procedendo con un'operazione di heapify sull'albero intero, si ottiene ancora una volta una catastrofica che rispetta il vincolo e si ricomincia il procedimento daccapo, fino ad ottenere l'array perfettamente ordinato, come mostrato di seguito:

Figura 27: Esempio di albero ordinato dopo $n - 1$ heap-sort

Di seguito si propone il listato dell'algoritmo **merge-sort**:

Algorithm 18 Heap-Sort

- 1: HEAP-SORT(A)
 - 2: BUILD-HEAP(A)
 - 3: **for** $i = \text{length}[A]$ **down to** 2
 - 4: **do** Exchange: $A[1] \leftrightarrow A[i]$
 - 5: heap-size[A] = heap-size[A] - 1
 - 6: HEAPIFY($A, 1$)
-

In questo caso il numero di iterazioni di ciclo for sono $n - 1$ e sono tutte operazioni in serie. Per valutare la complessità di tale algoritmo si deve tenere presente la seguente regola:

COMPLESSITÀ COMPLESSIVA DI UN ALGORITMO MULTI-BLOCCO

Quando un algoritmo è costituito da diversi blocchi di istruzioni, ognuno avente una propria distinta complessità esecutiva, prevale sempre la **complessità maggiore**.

In particolare, in questo caso, la complessità di build-heap è $\Theta(n)$, mentre la complessità nel secondo blocco è $O(n \cdot \log(n))$, essendo la complessità di heapify $O(\log(n))$ e dal momento che tale procedura si richiama $n - 1$ volte, applicandola su tutto l'albero e non su sottoalberi.

Pertanto si evince che la complessità dell'algoritmo **heap-sort** è al più log-lineare (ossia nel caso peggiore), ovvero

$$O(n \cdot \log(n))$$

mentre la complessità nel caso migliore è lineare, ovvero $\Theta(n)$. Richiamando il **teorema sugli algoritmi di ordinamento generale**, il quale afferma che un algoritmo di ordinamento di tipo generale, come **heap-sort**, è condannato ad avere una complessità nel caso peggiore almeno log-lineare e una complessità nel caso medio almeno log-lineare: questo significa che, siccome la complessità di **heap-sort** è al più log-lineare e tale risultato conferma che è almeno log-lineare, quindi, alla fine, la complessità dell'algoritmo **heap-sort** (nel caso peggiore e nel caso medio) è proprio **log-lineare**.

18 Marzo 2022

Si osservi che le operazioni che vengono svolte dall'algoritmo **heap-sort** sono **in serie**, pertanto devono essere valutate in maniera attenta per comprender la complessità dell'algoritmo stesso.

In particolare, è noto che, dal punto di vista della notazione Θ , se vi sono due blocchi di istruzioni che presentano complessità diverse, ciò che determina la complessità finale dell'algoritmo è la complessità peggiore e più pesante fra le due date. Ciò si può dimostrare facilmente, considerando, a tal proposito, due funzioni $f(n)$ e $g(n)$, allora

$$f(n) + g(n) \text{ e } f(n) \vee g(n)$$

in cui con $f(n) \vee g(n)$ è da intendersi il massimo tra $f(n)$ e $g(n)$, appartengono alla stessa classe Θ , in quanto, logicamente, si ha che

$$f(n) \vee g(n) \leq f(n) + g(n) \leq 2 \cdot [f(n) \vee g(n)]$$

In particolare, si osserva che l'algoritmo **heap-sort** procede dapprima realizzando una catasta tramite il sotto-algoritmo **build-heap**, la cui complessità è lineare; dopo questa procedura si ottiene un albero che è ordinato al contrario (con i valori maggiori sui primi nodi e i valori minori sugli ultimi nodi); pertanto si procede a sostituire il primo nodo con l'ultimo dei rimanenti ed eliminando progressivamente l'ultimo nodo e il collegamento con i rimanenti (tramite l'istruzione $\text{heap-size} = \text{heapsize} - 1$), procedendo ad ordinare l'albero intero (a partire dalla radice, a differenza della procedura *build-heap*, in cui *heapify* veniva applicata a sottoalberi vicini alle foglie), tramite la procedura *heapify*, ad ogni passaggio.

Pertanto la complessità dell'algoritmo **heap-sort** è **log-lineare**, che **asintoticamente** è ottima, come avvalorato dal teorema sulla complessità di un algoritmo di ordinamento generale.

4.8 Quick-sort

L'algoritmo di ordinamento **quick-sort** è un algoritmo che, come merge-sort, lavora su posizioni intermedie da 1 a n , impiegando gli indici p e r , e opera in maniera ricorsiva.

Tuttavia, mentre in merge-sort, ad ogni chiamata, l'insieme dei valori da ordinare si dimezzava fino a considerare un solo valore, in quick-sort tale divisione esatta in due parti non si verifica, in quanto le divisioni sono irregolari.

Il primo sotto-algoritmo adoperato da quick-sort prende il nome **partition** (partizione), il quale considera come argomenti l'array A di lunghezza n e due posizioni intermedie p e r , con il vincolo che $p < r$ (vincolo imposto da quick-sort), in cui inizialmente $p = 1$ e $r = n$, come mostrato di seguito:

p						r		
	2	8	7	2	3	5	6	4

Tale array verrà diviso in 4 aree, in una delle quali (l'ultima, generalmente) è presente un **perno** (o **pivot**): nella prima area saranno contenuti valori \leq del perno, nella seconda area saranno contenuti valori $>$ del perno, mentre nella terza area saranno contenuti valori che ancora non sono stati esaminati (e che quindi non è noto se siano maggiori o minori del perno): è chiaro che progressivamente le prime due aree andranno incrementandosi, mentre la terza area andrà via via riducendosi, fino a sparire con la fine dell'algoritmo. Pertanto, vi sono i confini delle due prime aree che sono mobili, chiamati i e j : il primo, i , che divide la prima e la seconda area, si muove in maniera irregolare; il secondo, j , che aumenta di un passo alla volta in modo meccanico.

Si osservi l'esecuzione di tale algoritmo: posto $p = 1$ e $r = n$, si ha $i = p - 1$ e $j = p$, come mostrato di seguito:

p						r		
	2	8	7	2	3	5	6	4
i	j							

e si procede a verificare se il primo elemento è più piccolo o più grande dell'ultimo elemento (ossia del perno), come nel caso seguente $2 \leq 4$; se questo è il caso, allora i si incrementa di 1 e si cambia $A[i] \leftrightarrow A[j]$ e si incrementa j di 1:

p						r		
	2	8	7	2	3	5	6	4
i	j							

Naturalmente in tale prima fase, lo scambio $A[i] \leftrightarrow A[j]$ è totalmente inutile, in quanto $i = j = 1$, ma avrà una validità significativa nel seguito.

Ora si confronta il j -esimo elemento con il pivot e si ha che $8 > 4$ e si incrementa j di uno, e così anche per l'iterazione successiva, in quanto $7 > 4$, giungendo alla configurazione seguente:

p	2		8	7	2	3	5	6	4	r
	i				j					

Tuttavia, quando $j = 4$, si ha che $2 \leq 4$, per cui si incrementa i di 1: tuttavia, adesso l'indice i punta ad una zona errata ($i = 2$ e quindi $A[i] = 8$), per cui si deve scambiare $A[i] \leftrightarrow A[j]$ ed incrementare j , ottenendo:

p	2	2		7	8	3	5	6	4	r
	i				j					

Confrontando il j -esimo elemento, ancora una volta $3 \leq 4$, per cui si incrementa i di 1 e si deve scambiare $A[i] \leftrightarrow A[j]$ ed incrementare j :

p	2	2	3		8	7	5	6	4	r
	i				j					

Le due ultime due iterazioni prevedono solo di incrementare j essendo $A[j]$ tutti valori maggiori del perno:

p	2	2	3		8	7	5	6		r
	i				j					

Ora da p a i vi sono valori più piccoli del pivot, mentre da $i + 1$ a j vi sono valori maggiori del perno. Ora è sufficiente scambiare l'ultima posizione con la posizione $i + 1$ per porre il perno nella sua posizione corretta:

p	2	2	3		4	7	5	6		r
	i				j					

Ecco che il **perno** è stato collocato proprio dove deve stare, all'interno dell'array: la sua posizione è definitiva e non verrà più cambiata, in quanto è preceduto da valori inferiori ed è seguito da valori maggiori: questa la chiave di funzione dell'algoritmo **quick-sort**.

Pertanto, ora, *partition* dovrà essere applicato da p ad i e da $i + 2$ a r , andando a collocare nella loro posizione esatta e definitiva tutti i perni rimanenti; ora che l'array è stato partizionato, di seguito si espone il listato dell'algoritmo **partition**:

Algorithm 19 Partition

```

1: PARTITION( $A, p, r$ )
2:  $x \leftarrow A[r]$ 
3:  $i = p - 1$ 
4: for  $j = p$  to  $r - 1$ 
5:   do if  $A[j] \leq x$ 
6:     then  $i = i + 1$ 
7:     Exchange:  $A[i] \leftrightarrow A[j]$ 
8:   Exchange:  $A[i + 1] \leftrightarrow A[r]$ 
9: return  $i + 1$ 

```

In cui, ovviaente $i + 1$ è la posizione definitiva del perno che viene restituita dal sotto-algoritmo *partition*.

Mentre di seguito si espone il listato dell'algoritmo **quick-sort**:

Algorithm 20 Quick-Sort

```

1: QUICK-SORT( $A, p, r$ )
2: if  $p < r$ 
3:   then  $q = \text{PARTITION}(A, p, r)$ 
4:   QUICK-SORT( $A, p, q - 1$ )
5:   QUICK-SORT( $A, q + 1, r$ )

```

In cui in q viene memorizzata la posizione definitiva del perno. Naturalmente ciò che permette di uscire dalla chiamata ricorsiva sono i tratti di lunghezza unitaria, per cui $p = r$.

Esempio: Si consideri l'array seguente:

3	4	1	2
---	---	---	---

Allora, nell'esecuzione dell'algoritmo **quick-sort**, si procede eseguendo dapprima la procedura *partition*, con $p = 1$ e $r = n$ e fissando come perno $A[r] = 2$.

Naturalmente, nelle prime due iterazioni, essendo $3 > 2$ e $4 > 2$, *partition* non opera alcun cambiamento all'array A , ed effettua esclusivamente l'incremento di j . Quando $j = 3$, allora si osserva che $1 \leq 2$, per cui si deve incrementare i di 1, scambiare $A[i] \leftrightarrow A[j]$ e incrementare j , ottenendo:

1	4	3	2
---	---	---	---

Arrivati all'ultima iterazione, bisogna eseguire lo scambio $A[i + 1] \leftrightarrow A[r]$, ottenendo:

1	2	3	4
---	---	---	---

ecco che il perno 2 si trova nella sua posizione $q = 2$ definitiva, dalla quale non verrà più spostato. Ora bisogna applicare quick-sort da $p = 1$ a $q - 1 = 1$ e da $q + 1 = 3$ a $r = 4$; naturalmente, nel primo caso, quick-sort non effettua alcuna operazione, visto che $p = r = 1$; nel secondo caso si dovrà applicare *partition* con $p = 3$ e $r = 4$, pertanto, essendo $3 \leq 4$, *partition* dovrà incrementare i di 1, eseguire lo scambio $A[i] \leftrightarrow A[j]$ (ma essendo $i = j$, tale scambio è superfluo) ed incrementare j di 1. Terminata l'unica iterazione del ciclo, *partition* eseguirà lo scambio $A[i + 1] \leftrightarrow A[r]$ (ma essendo $i + 1 = 4$ e $r = 4$, ancora una volta tale scambio è superfluo) per cui l'array diverrà

1	2	3	4
---	---	---	---

e *partition* terminerà, restituendo $q = 4$. Ancora una volta si dovrà applicare quick-sort, dalla posizione $p = 3$ e $r = 3$ e $p = 5$ e $r = 4$; naturalmente, in nessuno di tali casi $p < r$, per cui le chiamate ricorsive cessano, così come l'algoritmo quick-sort.

Osservazione: Per comprendere il numero di iterazioni che comporta l'utilizzo di quick-sort, non è possibile ragionare come con merge-sort, in cui venivano eseguite delle operazioni di divisione netta in due parti uguali ad ogni iterazione, in quanto con quick-sort non si ha la certezza che le partizioni su cui si andrà a lavorare siano proprio delle metà esatte.

Si consideri, a tal proposito, il caso di un ordinamento di n numeri che sono, in realtà, già ordinati:

1	2	3	4	...	$n - 1$		n
---	---	---	---	-----	---------	--	-----

Naturalmente, in questo, caso, il ciclo for della procedura *partition* eseguirà un numero di iterazioni pari a $n - 1$, ossia da 1 a $r - 1 = n - 1$: il perno si trovava in ultima posizione all'inizio del ciclo e taluna sarà la sua posizione definitiva; dopo il primo quick-sort, si dovranno applicare due quick-sort, uno da $p = 1$ a $r = q - 1 = n - 1$ e un secondo da $p = q + 1 = n + 1$ a $r = n$: è ovvio che dei due lavorerà solamente il primo, mentre il secondo prevederà solo un controllo iniziale; tuttavia, la prima *partition* prevede un ciclo for di $n - 2$ iterazioni, ossia da 1 a $r - 1 = n - 2$, e dopoiché si dovranno eseguire due nuovi quick-sort, il primo da $p = 1$ a $r = q - 1 = n - 2$ e il secondo da $p = q + 1 = n$ fino a $r = n - 1$; ancora una volta il primo prevede $n - 3$ iterazioni, mentre il secondo un solo controllo, e così via, fino ad arrivare ad un quick-sort da 1 iterazione e 1 controllo.

Naturalmente, la complessità data dal numero di controlli che devono essere eseguiti nei secondi quick-sort è $\Theta(n)$, mentre la somma di tutte le iterazioni dovute all'esecuzione dei primi quick-sort è data da

$$n - 1 + n - 2 + n - 3 + \dots + 1 = \frac{n \cdot (n + 1)}{2} \cong n^2$$

Ecco che allora si evince come la complessità di tale algoritmo, nel caso peggiore, è, appunto **quadratica**, ovvero $\Theta(n^2)$.

Tuttavia, la complessità dell'algoritmo nel caso generale (o caso medio) è **log-lineare**, che quindi rende tale algoritmo **asintoticamente** ottimo. Inoltre anche la **complessità empirica** è log-lineare, in quanto l'utilizzo di tale codice regolarmente conferma come la complessità quadratica si riscontri solamente in casi particolari, quando i dati in input sono strutturati in un certo modo.

4.9 Randomize Quick-Sort

Per comprendere la complessità dell'algoritmo **quick-sort**, nel caso generale, bisogna considerare una variante aleatoria di tale algoritmo, la quale prende il nome di **randomize quick-sort**.

La probabilità di un evento può essere definita come il grado di fiducia che il senso comune attribuisce al verificarsi dell'evento, valutato su una scala [0, 1]: tuttavia, la casualità dell'evento riguarda non tanto l'evento in sé, ma l'osservatore che, non potendo conoscere il risultato dell'evento stesso, lo reputa casuale.

Il concetto di probabilità, in senso teorico, può essere suddivisa in **tre macroaree di interesse**:

- probabilità combinatoria;
- probabilità empirica;
- probabilità soggettiva o neo-bayesiana.

La probabilità, infatti, è strettamente connessa al calcolo combinatorio: pertanto, la probabilità di un evento è definibile come il rapporto fra il numero di casi favorevoli e il numero di casi possibili (ossia gli eventi elementari), sempre nell'ipotesi che gli eventi elementari siano tutti equiprobabili (che è un'ipotesi decisamente condizionante per fornire significato alla formula seguente):

$$P(E) = \frac{\#\text{casi favorevoli}}{\#\text{casi possibili}}$$

per cui il grado di fiducia è proporzionale al numero di casi favorevoli.

Dal punto di vista empirico (o statistico o oggettivo, in quanto fondata sull'esperimento), invece, si ha che

$$P(E) = \lim_{n \rightarrow +\infty} \frac{\#\text{successi}}{n}$$

ovvero il limite metafisico quando il numero degli esperimenti diverge all'infinito del rapporto tra il numero dei successi e il numero degli esperimenti (è ovvio che sia un limite metafisico, perché affinché la mera valutazione statistica fornita da un numero finito di esperimenti si trasformi in probabilità, il numero degli esperimenti deve divenire infinito, cosa che è materialmente impraticabile). È chiaro che anche tale definizione è fortemente limitante, in quanto affinché essa sia affidabile, dovrebbe richiedere che tutti gli esperimenti che vengono eseguiti siano una copia dello stesso esperimento astratto, ossia che si ripetano sempre nelle stesse condizioni (**decidendo in maniera** soggettiva quali fattori siano rilevanti e, pertanto, debbano essere mantenuti costanti nel corso della sperimentazione): ecco, allora, che tale elemento soggettivo eguaglia quello della definizione combinatoria, nella quale si richiedeva l'equiprobabilità degli eventi elementari.

La probabilità soggettiva o neo-bayesiana (o definettiana) deve la sua concezione, naturalmente, al reverendo **Bayes**, mentre la sua diffusione e formulazione a **Bruno de Finetti**, il quale è riuscito ad analizzare gli *odds* tramite delle fondamenta teoriche precise, definendo la probabilità come la misura di quanto l'uomo è propenso a scommettere.

Si possono, pertanto, interpretare i tre differenti concetti di probabilità in **modalità telescopica**:

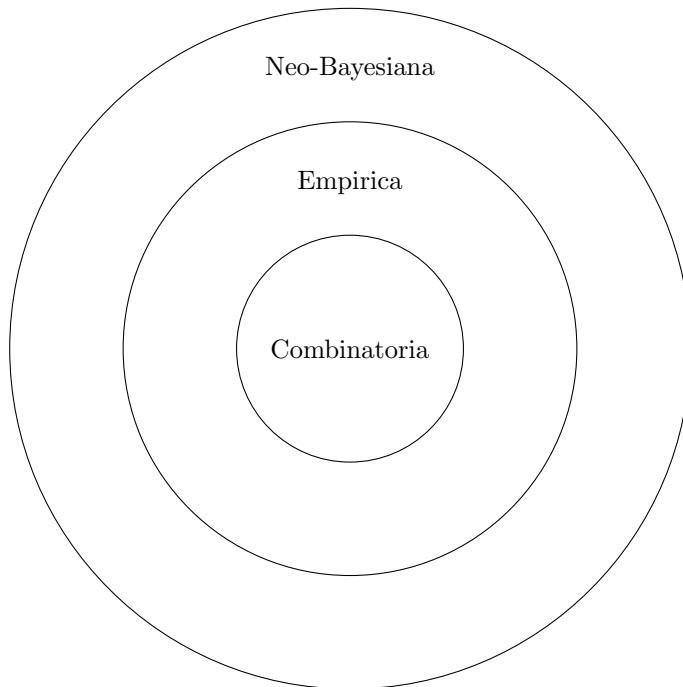


Figura 28: Interpretazione telescopica della probabilità

In cui è evidente come la probabilità combinatoria è contenuta nella probabilità statistica (o empirica) che sono contenute nella probabilità bayesiana, in quanto il suo campo di applicazione è enorme, notevolmente più esteso dei precedenti.

23 Marzo 2022

La definizione di probabilità più consona per una valutazione sperimentale è di tipo frequentista: la probabilità di un evento rappresenta il limite, per il numero degli esperimenti che tende all'infinito, del rapporto tra il numero di volte in cui tale evento si è verificato e il numero degli esperimenti effettuati (sempre nell'ipotesi che tali esperimenti ripetuti si siano svolti nelle medesime condizioni):

$$P(E) = \lim_{n \rightarrow +\infty} \frac{\#\text{successi}}{n}$$

in cui, ovviamente, essendo il numero degli esperimenti teoricamente infinito, la frequenza di successo dell'evento aleatorio E tende ad assestarsi su un valore che viene considerato la probabilità di E . Naturalmente, quello esposto è un limite metafisico in quanto un numero di esperimenti infinito è materialmente irrealizzabile.

È ovvio che, quindi, la probabilità che si verifichi l'evento E è

$$p(E) = \frac{\#\text{successi}}{n}$$

mentre la probabilità che non si verifichi E è

$$p(\bar{E}) = \frac{n - \#\text{successi}}{n}$$

Da ciò appare evidente come

$$P(\bar{E}) = 1 - P(E)$$

Considerando, ora, due eventi E e F , è nota che la probabilità dell'unione è proprio

$$P(E \vee F) = \frac{n_E + n_F - n_{E \wedge F}}{n} = p(E) + p(F) - p(E \wedge F)$$

Naturalmente, nel caso in cui E e F siano **incompatibili**, la **probabilità** diviene **additiva**, in quanto la probabilità che si verifichino entrambi i due eventi è nulla, allora

$$P(E \vee F) = p(E) + p(F)$$

Per quando concerne la **probabilità condizionata**, dati due eventi E e F , la probabilità che si verifichi E (**evento condizionato**) dopo che si è verificato F (**evento condizionante**), cosa che può essere vera o può essere una mera supposizione, si ottiene che:

$$P(E|F) = \frac{n_{E \wedge F}}{n_F} \cdot \frac{n}{n} = \frac{P(E \wedge F)}{p(F)}$$

ovvero il rapporto del numero degli esperimenti in cui si è verificato E e anche F e il numero di esperimenti in cui si è verificato F (altrimenti tale probabilità non sarebbe condizionata); tale risultato ha significato matematico solamente nell'ipotesi in cui l'evento condizionante abbia una probabilità strettamente positiva.

Nell'ipotesi in cui F ed E siano **indipendenti** l'uno rispetto all'altro (ovvero F non condiziona E e viceversa), è ovvio che

$$p(E|F) = P(E) \quad \text{e} \quad p(F|E) = P(F)$$

Ma ricordando la formula per il calcolo della probabilità di eventi condizionanti, si ottiene:

$$p(E|F) = P(E) = \frac{P(E \wedge F)}{p(F)} \quad \text{e} \quad p(F|E) = P(F) = \frac{P(E \wedge F)}{p(E)}$$

ed ecco che si è ottenuta la **relazione di indipendenza tra due eventi E e F** :

$$P(E \wedge F) = P(E) \cdot P(F)$$

per cui E non condiziona (o influenza) F se e solo se F non condiziona (o influenza) F .

Si consideri, ora, \mathcal{X} come **variabile** (o **quantità** o **qualità**) **aleatoria**, formata da uno spazio campionario (in questo caso finito) $\{x_1, \dots, x_n\}$ di **eventi** (o **determinazioni**) **lementari** i quali devono essere **incompatibili** fra di loro ed **esausitivi**: effettuando un esperimento, si deve sempre verificare uno e uno solo degli eventi elementari, in quanto l'esaurività prevede tutti gli eventi elementari possibili. Ciascuno di tali eventi elementari, inoltre, presenta una probabilità p_i , ottenendo l'insieme delle probabilità elementari $\{p_1, \dots, p_n\}$.

Naturalmente, siccome gli eventi elementari sono incompatibili fra di loro, la probabilità di un evento composto (o aggregato) è semplicemente la somma delle probabilità degli eventi elementari che lo compongono.

Nell'ipotesi in cui \mathcal{X} sia un **numero aleatorio**, ovvero $\mathcal{X} \in \Omega_x \subset \mathbb{R}$, allora ha senso calcolare il suo **valore medio** (o **valore atteso**), ovvero $E(\mathcal{X})$ (dall'inglese “expected value”, o speranza matematica).

Data una variabile aleatoria \mathcal{X} formata dagli eventi elementari seguenti:

VARIABLE ALEATORIA			
Eventi elementari	x_1	x_2	x_3
Prob. elementari	p_1	p_2	p_3

Pertanto, se x_1 figura n_1 volte, x_2 figura n_2 volte e x_3 figura n_3 volte, si ha che il valore medio di \mathcal{X} è

$$E(\mathcal{X}) = \frac{n_1 \cdot x_1 + n_2 \cdot x_2 + n_3 \cdot x_3}{n}$$

Allora è ovvio che

$$E(\mathcal{X}) = \sum_i^n p_i \cdot x_i$$

che risulta essere una definizione perfettamente consona della speranza matematica.

4.10 Teorema del trasferimento

Si consideri un numero aleatorio \mathcal{Y} (che deve essere un numero aleatorio, altrimenti non ha senso parlare della sua speranza matematica) che è funzione deterministica di una quantità (o qualità) aleatoria \mathcal{X} (la quale non deve essere necessariamente un numero aleatorio), ovvero

$$\mathcal{Y} = f(\mathcal{X})$$

tale per cui $f(a) = f(b) = f(d) = 0$, mentre $f(c) = 1$, in cui a, b, c, d sono valori di \mathcal{X} , aventi probabilità p_a, p_b, p_c, p_d . Naturalmente \mathcal{Y} è un **numero aleatorio di tipo binario**, di cui si vuole calcolare la speranza matematica; non impiegando il teorema del trasferimento, si deve procedere al **calcolo delle distribuzioni**, determinando le probabilità dei valori di \mathcal{Y} , cosa che è resa possibile in questo caso in quanto si opera su un insieme finito di valori, mentre dovrebbe essere eseguito mediante un espeditivo differente se si operasse sul continuo.

Allora appare evidente come

$$E(\mathcal{Y}) = \sum_y q_y \cdot y$$

in cui, per quanto che si è supposto, si ha $q_0 = p_a + p_b + p_d$ e $q_1 = p_c$, ottenendo

$$q_0 \cdot 0 + q_1 \cdot 1 = (p_a + p_b + p_d) \cdot 0 + p_c \cdot 1 = p_a \cdot f(a) + p_b \cdot f(b) + p_d \cdot f(d) + p_c \cdot f(c)$$

che corrisponde esattamente a

$$E(\mathcal{Y}) = \sum_x p_x \cdot f(x)$$

che riscritta in modo alternativo, permette di ottenere la formula del **teorema del trasferimento**:

$$E[f(\mathcal{X})] = \sum_x p_x \cdot f(x)$$

in cui non si richiede che \mathcal{X} sia di natura numerica, ma che soprattutto permette di conoscere $E[f(\mathcal{X})]$ senza dover procedere al calcolo delle distribuzioni di $f(\mathcal{X})$, ma lavorando unicamente con gli elementi di \mathcal{X} .

Applicando il teorema del trasferimento a due variabili aleatorie \mathcal{X} e \mathcal{Y} , si proceda al calcolo del valore atteso di $\mathcal{Z} = f(\mathcal{X}, \mathcal{Y}) = \mathcal{X} + \mathcal{Y}$, in cui \mathcal{Z} è una funzione deterministica della coppia \mathcal{X} e \mathcal{Y} . Si ottiene, quindi:

$$E[\mathcal{Z}] = E[\mathcal{X} + \mathcal{Y}] = \sum_x \sum_y p_{xy} \cdot f(x, y)$$

Ma siccome $f(\mathcal{X}, \mathcal{Y}) = \mathcal{X} + \mathcal{Y}$ per ipotesi, si ha che

$$\sum_x \sum_y p_{xy} \cdot f(x, y) = \sum_x \sum_y p_{xy} \cdot (x + y)$$

ma sfruttando la proprietà distributiva e commutativa (per cui l'ordine delle sommatorie può essere invertito) si ottiene

$$E[\mathcal{X} + \mathcal{Y}] = \sum_x \sum_y p_{xy} \cdot (x + y) = \sum_x \sum_y p_{xy} \cdot x + \sum_y \sum_x p_{xy} \cdot y$$

Ma siccome il primo termine non dipende dalla sommatoria in y e il secondo termine non dipende dalla sommatoria in x , si può scrivere:

$$\sum_x \sum_y p_{xy} \cdot x + \sum_y \sum_x p_{xy} \cdot y = \sum_x x \sum_y p_{xy} + \sum_y y \sum_x p_{xy}$$

Ma siccome

$$\sum_y p_{xy}$$

è la somma delle probabilità in cui x è fisso, mentre y è arbitrario, allora taluna è la probabilità che \mathcal{X} valga x e \mathcal{Y} sia arbitrario, ovvero la **probabilità marginale** di avere x , ossia p_x (in quanto il valore di y è totalmente ininfluente). Analogamente

$$\sum_x p_{xy}$$

è la somma delle probabilità in cui y è fisso, mentre x è arbitrario, allora taluna è la probabilità che \mathcal{Y} valga y e \mathcal{X} sia arbitrario, ovvero la **probabilità marginale** di avere y , ossia p_y (in quanto il valore di x è totalmente ininfluente). Pertanto la sommatoria di cui sopra corrisponde esattamente a

$$\sum_x x \sum_y p_{xy} + \sum_y y \sum_x p_{xy} = \sum_x x \cdot p_x + \sum_y y \cdot p_y = E(\mathcal{X}) + E(\mathcal{Y})$$

che permette di concludere che la **speranza matematica è additiva**: un risultato fondamentale, che è reso ancora più significativo in quanto valido in qualsiasi ipotesi di interdipendenza tra \mathcal{X} e \mathcal{Y} .

Esempio: Si consideri un **indicatore di successo** $I \in \{0, 1\}$, ossia un **numero aleatorio binario**, in cui $p_0 = 1 - p$ e $p_1 = p$ (per la formula del calcolo della probabilità dell'evento complementare). Allora è evidente come

$$E(I) = p_0 \cdot 0 + p_1 \cdot 1 = (1 - p) \cdot 0 + p \cdot 1 = p$$

Nell'ipotesi in cui \mathcal{Z} sia una funzione deterministica di un numero n di indicatori di successo, allora il numero di successi è dato da:

$$\mathcal{Z} = I_1 + I_2 + \dots + I_n$$

Anche se le modalità sperimentali garantiscono l'indipendenza tra gli indicatori di successo, in realtà non ve ne è bisogno, in quanto la speranza matematica è data dalla somma della speranza matematica di ciascun indicatore di successo, ed avendo visto che $E(I) = p_0 \cdot 0 + p_1 \cdot 1 = (1 - p) \cdot 0 + p \cdot 1 = p$ e che il numero di indicatori è n , è facile concludere che

$$E(\mathcal{Z}) = n \cdot p$$

Allora, senza impiegare il teorema del trasferimento, si può descrivere \mathcal{Z} tramite il calcolo delle distribuzioni, andando a determinare il numero di valori che \mathcal{Z} può assumere (0 successi, 1 successo, ..., $n - 1$ successi, n successi), come mostrato di seguito:

$$\mathcal{Z} = \sum_{i=0}^n \text{Prob}\{\mathcal{Z} = i\} \cdot i = \sum_i \binom{n}{i} p^i \cdot (1 - p)^{n-i} \cdot i = n \cdot p$$

Ovvero si è ottenuto lo stesso risultato ottenuto in precedenza, solamente che il processo esecutivo si basato sul calcolo della probabilità che \mathcal{Z} assuma il valore i -esimo, tenuto conto che \mathcal{Z} è una funzione deterministica di I_1, I_2, \dots, I_n , ossia di n indicatori di successo: è, in altre parole, un oggetto n -dimensionale.

Osservazione: Si consideri una stringa binaria di lunghezza $n = 5$ e di peso $i = 3$, come mostrato di seguito

$$\boxed{0 \ 1 \ 1 \ 0 \ 1}$$

e se ne consideri una seconda

$$\boxed{1 \ 1 \ 0 \ 1 \ 0}$$

che naturalmente è ottenuta permutando la stringa binaria precedente. Tali stringhe binarie sono da reputarsi come la sequenza degli esiti di n esperimenti binari, il cui risultato può essere o 0 o 1. Allora, in ciascuno di tali casi, la probabilità che si presenti una qualsiasi stringa di peso $i = 3$ e lunghezza $n = 5$ è pari

$$p^i \cdot (1 - p)^{n-i}$$

essendo gli esperimenti indipendenti e la probabilità moltiplicativa. Tale risultato è stato ottenuto in quanto si è assunto che la probabilità che, in ogni esperimento, figuri 1 è p , mentre la probabilità che figuri 0 è $1 - p$, per cui la probabilità di ottenere la stringa binaria richiesta è p elevato al numero dei successi i , moltiplicato per $1 - p$, elevato al numero degli insuccessi $n - p$.

Da ciò si evince che la probabilità di ottenere $i = 3$ successi, con qualunque stringa come quelle considerate, è data dalla somma delle probabilità di ottenere ciascuna stringa, essendo esse incompatibili (il risultato dei 5 esperimenti o è uno o è l'altro): pertanto, la probabilità che $\mathcal{Z} = i$ è proprio il prodotto tra il numero di stringhe aventi peso $i = 3$ e lunghezza $n = 5$ e la probabilità $p^i \cdot (1 - p)^{n-i}$, ossia

$$P\{\mathcal{Z} = i\} = \#\text{stringhe aventi peso } i=3 \text{ e lunghezza } n=3 \cdot p^i \cdot (1 - p)^{n-i}$$

in cui $\#\text{stringhe aventi peso } i=3 \text{ e lunghezza } n=3$ rappresenta il numero di stringhe di lunghezza $n = 5$ avente peso $i = 3$ che è, ovviamente:

$$\#\text{successi} = \binom{n}{i} = \frac{n!}{i! \cdot (n - i)!}$$

Osservazione: È noto che la speranza matematica binomiale che si verifichi la variabile aleatoria \mathcal{Z} è n volte la probabilità di successo, ovvero

$$E(\mathcal{Z}) = np = \mu_{\mathcal{Z}}$$

Tuttavia, per valutare propriamente un numero aleatorio, oltre alla speranza matematica necessario conoscere anche la **dispersione attorno al valore atteso** della variabile \mathcal{X} , ossia la **varianza**, che si potrebbe calcolare come

$$\text{var}(\mathcal{X}) = E[|\mathcal{X} - \mu|]$$

essendo la risposta più ovvia a tale problema. Tuttavia, volendo eliminare il valore assoluto si può procedere valutando il quadrato della differenza tra \mathcal{X} e la sua speranza matematica, ossia:

$$\text{var}(\mathcal{X}) = E[(\mathcal{X} - \mu)^2]$$

per cui la varianza della variabile aleatoria \mathcal{X} si definisce come **il valore atteso del quadrato della differenza tra \mathcal{X} e μ** . Al fine di ottenere, tuttavia, una valutazione lineare e non quadratica, molto spesso al posto della varianza si consiedera lo **scarto quadratico medio**

$$\text{sqm}(\mathcal{X}) = \sqrt{\text{var}(\mathcal{X})}$$

Naturalmente, però, dal teorema del trasferimento, è noto che la varianza si può scrivere come

$$\text{var}(\mathcal{X}) = \sum_x (x - \mu)^2 \cdot p_x$$

essendo $f(\mathcal{X}) = (\mathcal{X} - \mu)^2$ una funzione deterministica di \mathcal{X} ; sviluppando il quadrato descritto, tale risultato si può scrivere come

$$\sum_x x^2 \cdot p_x - 2\mu \cdot \sum_x x \cdot p_x + \mu^2 \cdot \sum_x p_x$$

ma siccome la somma di tutte le probabilità elementari è

$$\sum p_x = 1$$

e la speranza μ si può ravvisare nel calcolo intermedio

$$\sum x p_x = \mu$$

si ottiene che

$$\sum_x x^2 \cdot p_x - 2\mu \cdot \sum_x x \cdot p_x + \mu^2 \cdot \sum_x p_x = \sum_x x^2 \cdot p_x - \mu^2$$

Invocando ancora una volta il teorema del trasferimento, essendo $f(\mathcal{X}) = \mathcal{X}^2$ una funzione deterministica di \mathcal{X} , si ottiene che:

$$\text{var}(\mathcal{X}) = E[\mathcal{X}^2] - [E(\mathcal{X})]^2$$

Volendo calcolare la varianza della somma di due variabili aleatorie, si può procedere come segue:

$$\text{var}(\mathcal{X} + \mathcal{Y}) = E[(\mathcal{X} + \mathcal{Y})^2] - [E(\mathcal{X} + \mathcal{Y})]^2$$

tuttavia, in tale caso, è fondamentale considerare l'ipotesi di indipendenza seguente $p_{xy} = p_x \cdot q_y$, da cui si evince che la **varianza è additiva**:

$$\text{var}[\mathcal{X} + \mathcal{Y}] = \text{var}(\mathcal{X}) + \text{var}(\mathcal{Y})$$

Osservazione: Si osservi che, in questo caso, la varianza non è, in generale, additiva, come la speranza matematica: la varianza è additiva solamente nell'ipotesi di indipendenza tra \mathcal{X} e \mathcal{Y} . L'indipendenza è una condizione sufficiente, ma non necessaria, per l'additività: ci sono situazioni che prendono il nome di **non correlazione**, in cui vi è additività anche senza indipendenza.

Esempio: Considerando, ora, un indicatore di successo I , volendo calcolarne la varianza, applicando le formule note, si perviene al risultato seguente:

$$\text{var}(I) = E[I^2] - E[I]^2 = E[I^2] - p^2$$

siccome è noto che $E[I] = (1-p) \cdot 0 + p \cdot 1 = p$. Tuttavia, siccome $I^2 = I$, in quanto $0^2 = 0$ e $1^2 = 1$, allora si ha che $I^2 \in \{0, 1\}$, in cui, ancora una volta $p_0 = 1 - p$ e $p_1 = p$. Si ottiene, quindi, che

$$\text{var}(I) = E[I^2] - p^2 = E[I] - p^2 = p - p^2 = p(1 - p) = p \cdot q$$

essendo $q = 1 - p$ la probabilità di insuccesso. Pertanto, se si considera la variabile aleatoria \mathcal{Z} , la quale prevede di eseguire n esperimenti a ciascuno dei quali si associa un indicatore di successo I_i indipendente rispetto agli altri, allora la **varianza binomiale** è additiva (in quanto vi è indipendenza), per cui è la somma delle n varianze degli indicatori di successo (pari a $p \cdot q$), per cui si ottiene

$$\text{var}(\mathcal{Z}) = n \cdot p \cdot q$$

mentre il suo **valore atteso** è

$$E(\mathcal{Z}) = n \cdot p$$

Osservazione: Si consideri una variabile aleatoria \mathcal{X} e uno spazio campionario A tale che $|A| = k$, in cui la probabilità è uniforme, ovvero ciascuno degli eventi elementari presenta probabilità

$$P(E) = \frac{1}{|A|} = \frac{1}{k}, \quad \forall E \in A$$

Considerando un sottoinsieme $B \subset A$, tale che $|B| = h$, volendo calcolare la probabilità condizionata

$$P(X|B) = \text{Prob}\{\mathcal{X} = x | \mathcal{X} \in B\}$$

che ha significato, naturalmente, solamente quando $x \in B$. Alla luce di ciò si ha che

$$P(\mathcal{X}|B) = \frac{P(x \cap B)}{P(B)}$$

ma è chiaro che se $x \in B$, allora $P(x \cap B) = P(x)$, per cui si ottiene che

$$P(\mathcal{X}|B) = \frac{P(x \cap B)}{P(B)} = \frac{P(x)}{P(B)} = \frac{\frac{1}{k}}{h \cdot \frac{1}{k}} = \frac{1}{h}$$

in quanto la probabilità degli eventi in B si mantiene uniforme: tutti gli eventi di B hanno la stessa probabilità, uno rispetto all'altro.

Osservazione: Si consideri una variante aleatoria dell'algoritmo quick-sort, nella quale si deve procedere alla definizione della procedura $\text{RANDOM}(p,r)$, in cui si deve imporre che $p \leq r$, essendo essi due numeri interi, procedura capace di generare dei numeri che, tuttavia, saranno solamente pseudocasuali; naturalmente i numeri generati da p a r (estremi inclusi) sono $r - p + 1$, in cui ciascuno di tali valori presenta la probabilità

$$P(\text{num}) = \frac{1}{r - p + 1}$$

di essere generato ed estratto, nell'ipotesi in cui la **probabilità sia uniforme**, ossia identica per ciascun valore che significa che si ha totale casualità.

24 Marzo 2022

La serie armonica è così definita

$$\sum_{k=1}^{\infty} \frac{1}{k}$$

Tale serie, naturalmente, essendo a **termini positivi**, può divergere o convergere: è chiaro che una condizione necessaria affinché tale serie converga è che il termine generico della serie sia infinitesimo, ma non è una condizione sufficiente. Si consideri la **somma parziale** seguente:

$$\sum_{k=1}^{18} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \frac{1}{10} + \frac{1}{11} + \frac{1}{12} + \frac{1}{13} + \frac{1}{14} + \frac{1}{15} + \frac{1}{16} + \frac{1}{17} + \frac{1}{18}$$

Volendo ottenere una limitazione inferiore di tale somma, si può procedere al raggruppamento seguente

$$1 + \underbrace{\frac{1}{2}}_{\frac{1}{2}} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{>2 \cdot \frac{1}{4}} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{>4 \cdot \frac{1}{8}} + \underbrace{\frac{1}{9} + \frac{1}{10} + \frac{1}{11} + \frac{1}{12}}_{>8 \cdot \frac{1}{16}} + \underbrace{\frac{1}{13} + \frac{1}{14} + \frac{1}{15} + \frac{1}{16}}_{>8 \cdot \frac{1}{16}} + \frac{1}{17} + \frac{1}{18}$$

in cui sono state via via ottenute delle minorazioni sostituendo all'interno di ogni raggruppamento ogni frazione con il reciproco di una opportuna potenza di due, come mostrato di seguito:

$$\frac{1}{3} + \frac{1}{4} > \frac{1}{4} + \frac{1}{4} = 2^1 \cdot \frac{1}{4}$$

Il numero di addendi da sommare è dato proprio da potenze crescenti di 2. Il tratto interrotto e non raggruppato viene eliminato, dal momento che si sta cercando una limitazione inferiore, che è facile capire essere la seguente:

$$\sum_{k=1}^{18} \frac{1}{k} > 1 + \frac{1}{2} \cdot \lfloor \log_2(n) \rfloor$$

In cui è ovvio che il numero di $\frac{1}{2}$ che sono stati sommati è proprio pari alla parte intera inferiore del logaritmo in base 2 del numero totale n di rapporti sommati: nel caso seguente si ha

$$\lfloor \log_2(18) \rfloor = 4$$

in cui si ha disegualanza stetta a meno che $n \leq 2$. Ciò è sufficiente a dimostrare che la serie armonica diverge (nonostante il termine generico sia infinitesimo), dal momento che la limitazione inferiore contiene $\log_2(n)$ che in infinito, seppur sottoreale, per $n \rightarrow +\infty$.

Per voler ottenere una limitazione superiore, ora, si esegue il seguente raggruppamento:

$$1 + \underbrace{\frac{1}{2} + \frac{1}{3}}_{<2 \cdot \frac{1}{2}} + \underbrace{\frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7}}_{<4 \cdot \frac{1}{4}} + \underbrace{\frac{1}{8} + \frac{1}{9} + \frac{1}{10} + \frac{1}{11} + \frac{1}{12} + \frac{1}{13} + \frac{1}{14} + \frac{1}{15}}_{<8 \cdot \frac{1}{8}} + \underbrace{\frac{1}{16} + \frac{1}{17} + \frac{1}{18}}_{<16 \cdot \frac{1}{16}}$$

in cui sono state via via ottenute delle maggiorazioni sostituendo all'interno di ogni raggruppamento ogni frazione con il reciproco di una opportuna potenza di due, come mostrato di seguito:

$$\frac{1}{2} + \frac{1}{3} < \frac{1}{2} + \frac{1}{2} = 2^1 \cdot \frac{1}{2} = 1$$

Ancora una volta, il numero di addendi da sommare è dato proprio da potenze crescenti di 2. Il tratto interrotto e viene maggiorato fino al reciproco della potenza di due successiva, dal momento che si sta cercando una limitazione superiore, che è facile capire essere la seguente:

$$\sum_{k=1}^{18} \frac{1}{k} < 1 + \lfloor \log_2(n) \rfloor$$

Analogamente a quanto visto in precedenza, il numero di 1 che sono stati sommati è proprio pari alla parte intera inferiore del logaritmo in base 2 numero totale n di rapporti sommati: nel caso seguente si ha

$$\lfloor \log_2(18) \rfloor = 4$$

in cui si ha disuguaglianza stretta a meno che $n = 1$. Pertanto la serie armonica risulta così limitata:

$$1 + \frac{1}{2} \cdot \lfloor \log_2(n) \rfloor < \sum_{k=1}^n \frac{1}{k} < 1 + \lfloor \log_2(n) \rfloor$$

In cui le disuguaglianze sono strette almeno che, nel primo caso $n = 1 \vee n = 2$ e nel secondo caso $n = 1$. Volendo allora migliorare le limitazioni inferiori e superiori, si può scrivere:

$$1 + \frac{1}{2} \cdot \lfloor \log_2(n) \rfloor > 1 + \frac{1}{2} \cdot (\log_2(n) - 1) \geq \frac{1}{2} \cdot \log_2 n$$

inoltre si ha che

$$1 + \lfloor \log_2(n) \rfloor \leq 1 + \log_2(n) \leq \log_2(n) + \log_2(n) = 2 \cdot \log_2(n)$$

ed ecco quindi che si è ottenuta una nuova limitazione:

$$\frac{1}{2} \cdot \log_2(n) \leq \sum_{k=1}^n \frac{1}{k} \leq 2 \cdot \log_2(n)$$

che permette di concludere che

$$\sum_{k=1}^n \frac{1}{k} = \Theta(\log_2(n))$$

anche se la base del logaritmo è totalmente irrilevante per la notazione Θ .

Osservazione: La procedura **partition** veniva richiamata ricorsivamente dall'algoritmo **quick-sort** e permetteva di collocare progressivamente ogni pivot nella loro posizione definitiva (in quanto seguiti da valori maggiori e preceduti da valori minori); tuttavia, nel caso peggiore, è già noto che la complessità di quick-sort è quadratica. All'interno di partition vi è un ciclo for che presenta un numero di iterazioni pari a $r - 1 - p + 1 = r - p$; tuttavia, partition viene richiamato molteplici volte all'interno di quick-sort, per cui al fine di determinare la complessità dell'algoritmo è necessario sommare tutte le iterazioni che vengono effettivamente eseguite.

Si può osservare, però, che il numero di iterazioni che viene eseguito da partition è pari al **#confronti con il pivot**, in quanto in ogni iterazione vi è, per ogni valore, uno e un solo confronto con esso: contare il numero totale delle iterazioni for all'interno di quick-sort equivale a contare il numero totale di confronti con il pivot che vengono effettuati nel corso di tutto l'algoritmo nell'uso di ogni partition.

Oltre a ciò, tuttavia, è necessario considerare un'ulteriore componente di complessità; a tal proposito, si osservi il listato dell'algoritmo quick-sort:

Algorithm 21 Quick-Sort

```

1: QUICK-SORT( $A, p, r$ )
2: if  $p < r$ 
3:   then  $q = \text{PARTITION}(A, p, r)$ 
4:   QUICK-SORT( $A, p, q - 1$ )
5:   QUICK-SORT( $A, q + 1, r$ )

```

Come si vede, naturalmente, oltre ai confronti col pivot all'interno di partition, si devono considerare anche i confronti tra p e r che “vanno a vuoto”, ossia quando $p = r$ e quando $p = r + 1$; il primo caso si ha quando il perno viene collocato in penultima posizione al termine di partition e quindi si applica nuovamente partition tra $p = r$ e $r = r$, mentre il secondo caso si ha quando il perno viene collocato in ultima posizione (in quanto era il maggiore tra tutti i valori fin dall'inizio) e quindi applicando nuovamente partition si ha che $p = r + 1$ e $r = r$. In ambedue i casi si deve fare uno sforzo $\Theta(1)$ per avere la certezza di aver finito la procedura: il numero di confronti negativi che si verificano quando $p = r$ può essere al massimo $\leq n$, mentre nel secondo caso il numero dei confronti negativi che si possono verificare quando $p = q + 1$ è al più $\leq n + 1$, per cui il numero di negativi che si devono valutare è sicuramente al più $\leq 2n + 1$ (ossia è **linearmente controllato**). Pertanto, si ha che il numero di iterazioni di quick-sort (ossia la sua complessità) è dato da

$$\#confronti col pivot + \#confronti negativi$$

ma è noto che, dal punto di vista della notazione Θ , tra le due quantità prevale la peggiore, che è naturalmente la prima, in quanto è ovvio che anche solo per scrivere i dati lo sforzo è almeno lineare, per cui il numero di iterazioni di quick-sort è valutabile semplicemente tramite il conteggio del numero di confronti con il pivot.

È noto che nel caso peggiore (come quando i numeri sono già ordinati), quick-sort presenta una complessità quadratica, mentre negli altri casi è importante valutare la scelta della posizione di collocamento del pivot: a tal proposito si deve introdurre una variante aleatoria della procedura partition, che prende il nome di **Randomized-Partition**, nella quale il perno viene scelto arbitrariamente tra uno dei valori compresi tra p e r (estremi inclusi):

Algorithm 22 Randomized-Partition

-
- 1: RANDOMIZED-PARTITION(A,p,r)
 - 2: $i = \text{RANDOM}(p,r)$
 - 3: $A[i] \leftrightarrow A[r]$
 - 4: PARTITION(p,r)
-

In cui, naturalmente, la probabilità di estrarre un numero tra p e r è casuale (e quindi uniforme)

$$P(\text{num}) = \frac{1}{r - p + 1}$$

Naturalmente l'istruzione $A[i] \leftrightarrow A[r]$ garantisce che il perno si trovi in una posizione perfettamente casuale, che quindi slega l'algoritmo dal vincolo di porre il pivot nell'ultima posizione.

La variante aleatoria di quick-sort è presto conclusa:

Algorithm 23 Randomized-Quick-Sort

-
- 1: RANDOMIZED-QUICK-SORT(A,p,r)
 - 2: **if** $p < r$
3: **then** $q = \text{RANDOMIZED-PARTITION}(A,p,r)$
 - 4: RANDOMIZED-QUICK-SORT($A,p,q - 1$)
 - 5: RANDOMIZED-QUICK-SORT($A,q + 1,r$)
-

Ecco che allora, in questo caso, il numero di confronti con il pivot non è determinabile a priori, ma costituisce una **variabile aleatoria** \mathcal{X} , di cui si può formulare una previsione del suo valore, determinandone la speranza matematica.

Per effettuare tale valutazione, per semplicità si supponga, senza perdita di generalità, che gli n numeri da ordinare siano tutti diversi fra di loro tale che, una volta operato l'ordinamento, si ottiene

$$z_1 < z_2 < z_3 < \dots < z_n$$

Alla luce di ciò si considerino due valori arbitrari z_i e z_j , con $i \neq j$: è ovvio che è possibile che tra di essi non avvenga alcun confronto (come nel caso in cui $z_i \leq$ perno $\leq z_j$), oppure che vengano confrontati una e una sola volta, in quanto se tale confronto avviene, ciò significa che uno dei due è un perno e dopo che il perno è stato collocato nella sua posizione definitiva, esso non viene più sottoposto ad alcuna operazione di confronto.

Allora si ha che

$$\boxed{\mathcal{X} = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \mathcal{X}_{i,j}}$$

in cui si è supposto $i < j$, mentre $\mathcal{X}_{i,j}$ è un **numero aleatorio** che quantifica la probabilità che z_i e z_j vengono confrontati nel corso dell'algoritmo: in altre parole $\mathcal{X}_{i,j}$ rappresenta un indicatore di successo (o un **bit aleatorio**, che assume valore 1 quando avviene il confronto, 0 quando non avviene), che valuta il confronto di due termini z_i e z_j .

Volendo conoscere la speranza matematica di \mathcal{X} , si può sfruttare il fatto che la speranza matematica sia additiva: la speranza matematica di un numero finito di addendi è la somma delle speranze

matematiche; ciò permette di concludere che

$$E(\mathcal{X}) = E \left(\sum_{i=1}^{n-1} \sum_{j=i+1}^n \mathcal{X}_{i,j} \right) = \sum_{i=1}^{n-1} \sum_{j=i+1}^n E(\mathcal{X}_{i,j})$$

pertanto è sufficiente calcolare la speranza matematica di $\mathcal{X}_{i,j}$ ed effettuarne la somma. Inoltre, è ovvio capire che la speranza matematica del bit aleatorio $\mathcal{X}_{i,j}$ è proprio la probabilità che z_i e z_j vengano confrontati, quindi

$$E(\mathcal{X}_{i,j}) = 0 \cdot \text{Prob}(\text{insuccesso}) + 1 \cdot \text{Prob}(\text{successo}) = \text{Prob}\{z_i \text{ cfr } z_j\}$$

Nell'ipotesi in cui né z_i né z_j siano un perno e che il perno scelto sia $z_i \leq$ perno $\leq z_j$, allora z_i e z_j non verranno mai confrontati fra di loro, naturalmente, in quanto il perno, alla fine della procedura partition, si posizionerà tra z_i e z_j e partition si applicherà nei due tratti distinti. Se ne deduce che l'unica possibilità di un confronto tra z_i e z_j è che o l'uno o l'altro siano il primo pivot scelto nell'intervallo $[z_i, z_j]$; la probabilità che ne deriva è data dalla somma di due probabilità, come mostrato di seguito:

$$\text{Prob}\{z_i \text{ cfr } z_j\} = \text{Prob}\{z_i \text{ o } z_j \text{ è il primo perno estratto da } Z_{i,j}\}$$

in cui $Z_{i,j}$ è l'insieme di tutti i numeri compresi tra z_i e z_j (estremi inclusi): ma ciò fa capire come taluna sia una **probabilità condizionata** dal fatto che il valore estratto appartenga all'insieme $Z_{i,j}$, ovvero:

$$\text{Prob}\{z_i \text{ cfr } z_j \mid Z_{i,j}\}$$

Ricordando che la probabilità uniforme su insieme di valori si mantiene uniforme quando si condiziona su un sottoinsieme dell'insieme di partenza (come mostrato in precedenza sull'insieme B dell'esempio precedente). Ecco che allora la probabilità cercata è

$$E(\mathcal{X}_{i,j}) = \text{Prob}\{z_i \text{ cfr } z_j\} = \frac{2}{j - i + 1}$$

in quanto due sono gli eventi da sommare (ossia z_i e z_j , che hanno probabilità uniforme) e $|Z_{i,j}| = j - i + 1$. Pertanto si ottiene che

$$E[\mathcal{X}] = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{2}{j - i + 1}$$

Volendo semplificare l'espressione, si opera una sostituzione con $k = j - 1$, sicché se j partiva da $i + 1$ e arrivava a n , k parte da 1 e giunge fino a $n - i$, allora si può ottenere:

$$E[\mathcal{X}] = \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k + 1}$$

Volendo ottenere una maggiorazione di tale aspettativa matematica, si può fare in modo che la seconda sommatoria giunga fino a n e l'1 del denominatore venga eliminato, ottenendo:

$$E[\mathcal{X}] < 2 \cdot \sum_{i=1}^{n-1} \sum_{k=1}^n \frac{1}{k}$$

in cui la seconda sommatoria è proprio la serie armonica precedentemente calcolata e di cui si era visto che

$$\sum_{k=1}^n \frac{1}{k} = \Theta(\log_2(n))$$

Per cui si può scrivere:

$$E[\mathcal{X}] < 2 \cdot \sum_{i=1}^{n-1} \sum_{k=1}^n \frac{1}{k} < 2 \cdot \sum_{i=1}^n \sum_{k=1}^n \frac{1}{k} = 2 \cdot \sum_{i=1}^n \Theta(\log_2(n)) = 2\Theta(n \cdot \log_2(n)) = \Theta(n \cdot \log_2(n))$$

Naturalmente, trattandosi di una maggiorazione, si ottiene che:

$$E(\mathcal{X}) = O(n \cdot \log(n))$$

e ricordando il teorema sulla complessità di un algoritmo di ordinamento generale, si ha anche che $E(\mathcal{X}) = \Omega(n \cdot \log(n))$, per cui si ottiene che:

$$E(\mathcal{X}) = \Theta(n \cdot \log(n))$$

Per cui si ottiene che la complessità di randomized-quick-sort **tipicamente è log-lineare**: più specificatamente è noto che nel caso peggiore la complessità è quadratica, mentre nel caso medio (o tipico) e nel caso migliore, la complessità è **log-lineare** (quindi si ha che tipicamente l'algoritmo opera molto più vicino al caso migliore).

Osservazione: Si osservi che gli algoritmi fino ad adesso esposti non alterano la loro esecuzione a seconda della natura o grandezza dei numeri da ordinare.

Pertanto, gli algoritmi di ordinamento generale (per **confronti**) presentano una complessità nel caso peggiore è almeno log-lineare:

$$T_{\text{worst}}(n) = \Omega(n \cdot \log(n))$$

mentre nel caso medio si ha una complessità, ancora una volta, almeno log-lineare:

$$T_{\text{avarage}}(n) = \Omega(n \cdot \log(n))$$

È chiaro che sapendo il secondo risultato, naturalmente si ottiene anche il primo.

Osservazione: Per dimostrare che un algoritmo di ordinamento per confronti presenta una complessità almeno log-lineare, si osservi che un algoritmo per confronti, al fine di procedere all'ordinamento di n valori distinti, procede alla costruzione di un albero binario avente $n!$ foglie, in cui si trovano proprio gli ordinamenti possibili di n valori distinti.

25 Marzo 2022

In un algoritmo di ordinamento generale l'elemento fondamentale è dato dai confronti, i quali vanno a costituire un albero binario, sulle cui foglie sono contenuti tutti i possibili ordinamenti dei valori di partenza: in particolare, ogni algoritmo di ordinamento di tipo generale presenta una complessità che è dettata unicamente dai confronti.

È ovvio che, presi in considerazione n numeri distinti, il numero di ordinamenti possibili è dato da $n!$: in un albero binario, qualunque sia l'algoritmo di ordinamento generale considerato (efficiente o meno), tali ordinamenti sarebbero tutti collocati sulle foglie: ecco che il modo per osservare la **complessità nel caso peggiore** è quello di determinare l'**altezza dell'albero** (da intendersi come altezza massima, ovvero la distanza tra la radice e la foglia più lontana); inoltre la **complessità media** è data dalla **media delle distanze** delle foglie dalla radice (ovvero la somma delle distanze dalla radice di tutte le foglie diviso per $n!$, con $n!$ numero di foglie).

La soluzione più auspicabile al fine di ottimizzare sia la complessità media che la complessità massima, evitando che vi sia una differenza di distanza superiore a 1 è quella di andare a costruire (tramite un algoritmo opportunamente definito) un **albero binario completo**, riempiendo completamente tutti i livelli: taluna è la situazione più favorevole, in quanto la differenza tra l'altezza massima e quella minima è proprio 1 (per cui dal punto di vista della notazione Θ non cambia nulla).

Pertanto, sapendo che l'altezza di un albero binario completo è data dal logaritmo del numero delle foglie, si deduce che la complessità nel caso peggiore di un algoritmo di ordinamento che è in grado di produrre dei confronti che si diramano su un albero binario completo è

$$T_{\text{worse}}(n) = \Omega(\log_2(n!))$$

Questo, naturalmente, quando vi è un algoritmo capace di costruire un albero binario completo, altrimenti la situazione (così come la complessità), non può che peggiorare (e tale complessità è la stessa anche nel caso medio).

Osservazione: Si osservi che nel caso dell'algoritmo quick-sort, presa una qualsiasi permutazione di n numeri in ingresso, essa presenterà la stessa probabilità di presentarsi rispetto alle altre nell'ordinamento, la quale è pari a

$$\frac{1}{n!}$$

Naturalmente, l'esecuzione dell'algoritmo quick-sort presenta un elemento di casualità, in quanto dipende dall'estrazione di una possibile permutazione delle $n!$ n -uple in ingresso. Tuttavia, dal punto di vista di quick-sort, estrarre casualmente una permutazione di n numeri o estrarre di volta in volta in maniera casuale un perno, rappresentano la stessa cosa (cosa che si può facilmente dimostrare) in quanto non si sta privilegiando una configurazione rispetto ad un'altra: la complessità media di quick-sort e di randomized-quicksort permane sempre log-lineare.

È noto che la complessità nel caso medio di un algoritmo di ordinamento basato sui confronti è almeno $\log_2(n!)$ (con n numero di oggetti da ordinare); per capire l'entità del fattoriale di un numero n è opportuno richiamare la **formula di Stirling**:

$$n! \cong \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n = \text{St}(n)$$

in cui $\text{St}(n)$ è un'abbreviazione di formula di Stirling, la quale approssima il fattoriale di un numero n . Tale approssimazione non controlla certamente l'errore assoluto, ma controlla l'errore relativo (o percentuale), in quanto

$$\lim_{n \rightarrow +\infty} \frac{n!}{\text{St}(n)} = 1$$

In particolare si ha che il rapporto di cui sopra è pari alla somma di 1 e di una funzione infinitesima, come mostrato di seguito

$$\frac{n!}{\text{St}(n)} = 1 + \phi(n)$$

in cui $\phi(n)$ è una funzione infinitesima per $n \rightarrow \infty$, come quella mostrata di seguito:

$$1 < \frac{n!}{\text{St}(n)} < 1 + \frac{1}{11n}$$

in cui si può osservare come, in generale, $\text{St}(n) < n!$. Per questo si preferisce scrivere, dal punto di vista formale:

$$n! = \text{St}(n) \cdot \left[1 + O\left(\frac{1}{n}\right) \right]$$

In cui dal punto di vista della notazione O-grande, ciò significa che

$$1 + O\left(\frac{1}{n}\right) := \exists c \geq 0 : 1 + c \cdot \frac{1}{n} \leq 1 + \frac{1}{n}$$

in cui è noto, come mostrato nel precedente esempio, che $c = \frac{1}{11}$.

Tuttavia, è noto che il rapporto tra $n!$ e $\text{St}(n)$ per $n \rightarrow +\infty$ è uguale a 1, quindi hanno lo stesso ordine di infinito e quindi appartengono alla stessa classe Θ , per cui

$$\Theta(n!) = \Theta(\text{St}(n))$$

Ma è facile dimostrare che se due infiniti stanno nella stessa classe Θ , anche i loro logaritmi appartengono alla stessa classe Θ . Infatti, se è noto che

$$\exists \bar{x}, c_1 \geq 0, c_2 \geq 0 : c_1 \cdot g(x) \leq f(x) \leq c_2 \cdot g(x), \quad \forall x \geq \bar{x}$$

Ma giacché per le proprietà dei logaritmi si può scrivere

$$g(x) = a^{\log_a(g(x))}$$

essendo ambedue le funzioni definitivamente positive, allora si ha che:

$$\exists \bar{x}, c_1 \geq 0, c_2 \geq 0 : a^{\log_a(c_1)} \cdot a^{\log_a(g(x))} \leq a^{\log_a(f(x))} \leq a^{\log_a(c_2)} \cdot a^{\log_a(g(x))}, \quad \forall x \geq \bar{x}$$

E potendo ora trasferire la diseguaglianza agli esponenti, essendo tutte potenze con la stessa base si ottiene

$$\exists \bar{x}, c_1 \geq 0, c_2 \geq 0 : \log_a(c_1) + \log_a(g(x)) \leq \log_a(f(x)) \leq \log_a(c_2) + \log_a(g(x)), \quad \forall x \geq \bar{x}$$

Volendo ottenere una minorazione nel primo caso e una maggiorazione nel secondo è possibile sostituire $\log_a(c_1)$ con $-\frac{1}{2} \cdot \log_a(g(x))$ (essendo una quantità negativa) e $\log_a(c_2)$ con $\log_a(g(x))$ (essendo una funzione infinita), ottenendo

$$\exists \bar{x} : \frac{1}{2} \cdot \log_a(g(x)) \leq \log_a(f(x)) \leq 2 \cdot \log_a(g(x)), \quad \forall x \geq \bar{x}$$

Da cui si evince che:

$$\Theta(\log(n!)) = \Theta(\log(\text{St}(n)))$$

In cui si può calcolare che

$$\log(\text{St}(n)) = \log(\sqrt{2\pi n}) + n \cdot \log(n) - n \cdot \log(e)$$

Al fine di ottenere una maggiorazione, si può eliminare il termine negativo e , siccome, $n \cdot \log(n) \geq \log(\sqrt{2\pi n})$ (nel senso che il primo è un infinito di ordine superiore al secondo), si può scrivere

$$\log(\text{St}(n)) \leq n \cdot \log(n) + n \cdot \log(n) + 0 = 2n \cdot \log(n)$$

Al fine di ottenere una minorazione, invece, si elimina il termine $\log(\sqrt{2\pi n})$ e si sottrae al rimanente $\frac{1}{2}n \cdot \log(n)$, ossia un infinito log-lineare, in luogo di $n \cdot \log(e)$ che è solo un infinito lineare, ottenendo

$$\log(\text{St}(n)) \geq n \cdot \log(n) - n \cdot \log(e) \geq n \cdot \log(n) - \frac{1}{2}n \cdot \log(n) = \frac{1}{2}n \cdot \log(n)$$

Da cui si evince che l'approssimazione di Stirling appartiene ad una classe log-lineare, ed appartenendo anche alla classe del logaritmo del fattoriale di n se ne deduce che

$$\boxed{\Theta(\log(n!)) = \Theta(n \cdot \log(n))}$$

Pertanto, si ha che la complessità di un algoritmo di ordinamento generale (ossia per confronti), nel caso medio e nel caso peggiore è $\Omega(n \cdot \log(n))$.

5 Ordinamento di strutture dati specifiche

Gli algoritmi fin d'ora considerati sono totalmente indipendenti dalle strutture e dalla natura dei dati da ordinare, mentre vi sono alcuni algoritmi che funzionano in maniera eccelsa quando in ingresso vi sono degli specifici pattern di valori.

Pertanto, in tal senso, non si contraddice il teorema appena dimostrato che afferma che un algoritmo di ordinamento generale (basato sui confronti) presenta una complessità media e peggiore almeno log-lineare, in quanto gli algoritmi di seguito esposti **non sono basati sui confronti**, in quanto **strettamente correlati alla natura dei dati** (e non al loro ordinamento iniziale).

5.1 Counting-Sort

Si considerino $n = 200000$ **valori interi positivi** da ordinare, i quali, tuttavia, sono tutti valori compresi tra 0 e k , con $k \ll n$: è chiaro che i valori da ordinare saranno **duplicati molteplici volte**.

È possibile, naturalmente, effettuare delle generalizzazioni dell'algoritmo: se i numeri da ordinare andassero da r a $r+k$, basterebbe sottrarre a ciascuno r e si otterrebbe la configurazione di partenza (con valori compresi tra 0 e k); si potrebbe, inoltre, addattare l'algoritmo all'ordinamento non solo di numeri interi, ma anche di numeri razionali o irrazionali, codificandoli opportunamente.

Si propone di seguito il listato dell'algoritmo, il quale non rappresenta un ordinamento in loco, ma necessita di un array ausiliario B per effettuare l'ordinamento:

Algorithm 24 Counting-Sort

```
1: COUNTING-SORT( $A, B, k$ )
2: for  $i = 0$  to  $k$ 
3:   do  $C[i] = 0$ 
4: for  $j = 1$  to  $n$ 
5:   do  $C[A[j]] = C[A[j]] + 1$ 
6: for  $i = 1$  to  $k$ 
7:   do  $C[i] = C[i] + C[i - 1]$ 
```

In cui il numero di iterazione del primo ciclo è $k + 1$, nel secondo ciclo è n , mentre nel terzo vi sono k iterazioni: in totale si hanno $n + 2k + 1$ iterazioni.

All'inizio dell'algoritmo, con il primo ciclo for, si costruisce l'array C come segue:

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 3 \end{bmatrix}$$

e si supponga che l'array A da ordinare sia

$$A = \begin{bmatrix} 0 & 3 & 0 & 3 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}$$

Nel secondo ciclo for si deve procedere ad incrementare i valori all'interno dell'array C , come mostrato di seguito:

$$C = \begin{bmatrix} 2 & 1 & 0 & 2 \\ 0 & 1 & 2 & 3 \end{bmatrix}$$

in cui semplicemente si osserva come all'interno di tale array sono contenuti i conteggi di ogni valore presente in A , corrispondenti alle posizioni dell'array C : la cella $C[0]$ memorizza il numero di 0 all'interno dell'array A , la cella $C[1]$ memorizza il numero di 1 all'interno dell'array A e così via.

Nell'ultimo ciclo for si procede ad incrementare ogni elemento di ogni posizione con il valore della posizione precedente, ottenendo

$$C = \begin{bmatrix} 2 & 3 & 3 & 5 \\ 0 & 1 & 2 & 3 \end{bmatrix}$$

in cui in ogni posizione è contenuto il conteggio dei numeri che nell'array A sono minore (o uguali) della posizione dell'array C : nella cella $C[0]$ viene memorizzato il numero di valori che nell'array A sono minori (o uguali) di 0, nella cella $C[1]$ viene memorizzato il numero di valori che nell'array

A sono minori (o uguali) di 1, e così via. Per giungere a tale configurazione è stato effettuato uno sforzo di $n + 2k + 1$ (corrispondente ad una complessità $\Theta(n + k)$); si propone, adesso, la continuazione dell'algoritmo precedente:

Algorithm 25 Counting-Sort

```

1: COUNTING-SORT( $A, B, k$ )
2: for  $i = 0$  to  $k$ 
3:   do  $C[i] = 0$ 
4:   for  $j = 1$  to  $n$ 
5:     do  $C[A[j]] = C[A[j]] + 1$ 
6:   for  $i = 1$  to  $k$ 
7:     do  $C[i] = C[i] + C[i - 1]$ 
8:   for  $j = n$  down to 1
9:     do  $B[C[A[j]]] = A[j]$ 
10:     $C[A[j]] = C[A[j]] - 1$ 
```

Con l'ultimo ciclo introdotto, il numero di iterazioni complessive dell'algoritmo diviene $2n + 2k + 1$, che corrisponde alla classe $\Theta(n + k)$.

L'ultimo ciclo (che va al contrario, anche se non cambierebbe nulla), semplicemente, va a collocare i valori di A correttamente all'interno di B ; avendo a disposizione gli array seguenti:

$$A = \boxed{0 \ 3 \ 0 \ 3 \ 1} \\ \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array}$$

$$C = \boxed{2 \ 3 \ 3 \ 5} \\ \begin{array}{ccccc} 0 & 1 & 2 & 3 \end{array}$$

$$B = \boxed{0 \ 0 \ 0 \ 0 \ 0} \\ \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array}$$

Si procede come segue: posto $j = n$, l'algoritmo valuta che vi sono $C[A[j]]$ valori che sono minori o uguali a $A[j]$, per cui basta porre in $B[C[A[j]]]$ il valore $A[j]$ cercato e decrementare il conteggio in $C[A[j]]$, ottenendo:

$$A = \boxed{0 \ 3 \ 0 \ 3 \ 1} \\ \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array}$$

$$C = \boxed{2 \ 2 \ 3 \ 5} \\ \begin{array}{ccccc} 0 & 1 & 2 & 3 \end{array}$$

$$B = \boxed{0 \ 0 \ 1 \ 0 \ 0} \\ \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array}$$

Procedendo con $j = n - 1$ si ottiene

$$A = \boxed{0 \ 3 \ 0 \ 3 \ 1} \\ \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array}$$

$$C = \boxed{2 \ 2 \ 3 \ 4} \\ \begin{array}{ccccc} 0 & 1 & 2 & 3 \end{array}$$

$$B = \boxed{0 \ 0 \ 1 \ 0 \ 3} \\ \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array}$$

Procedendo con $j = n - 2$ si ottiene

$$A = \boxed{0 \ 3 \ 0 \ 3 \ 1} \\ \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array}$$

$$C = \boxed{1 \ 2 \ 3 \ 4} \\ \begin{array}{ccccc} 0 & 1 & 2 & 3 \end{array}$$

$$B = \boxed{0 \ 0 \ 1 \ 0 \ 3} \\ \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array}$$

Procedendo con $j = n - 3$ si ottiene

$$A = \boxed{0 \ 3 \ 0 \ 3 \ 1} \\ \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array}$$

$$C = \begin{array}{|c|c|c|c|} \hline 1 & 2 & 3 & 3 \\ \hline 0 & 1 & 2 & 3 \\ \hline \end{array}$$

$$B = \begin{array}{|c|c|c|c|c|} \hline 0 & 0 & 1 & 3 & 3 \\ \hline 1 & 2 & 3 & 4 & 5 \\ \hline \end{array}$$

Procedendo con $j = n - 4$ si ottiene

$$A = \begin{array}{|c|c|c|c|c|} \hline 0 & 3 & 0 & 3 & 1 \\ \hline 1 & 2 & 3 & 4 & 5 \\ \hline \end{array}$$

$$C = \begin{array}{|c|c|c|c|} \hline 0 & 2 & 3 & 3 \\ \hline 0 & 1 & 2 & 3 \\ \hline \end{array}$$

$$B = \begin{array}{|c|c|c|c|c|} \hline 0 & 0 & 1 & 3 & 3 \\ \hline 1 & 2 & 3 & 4 & 5 \\ \hline \end{array}$$

Per valutare la complessità globale dell'algoritmo, si parte dalla supposizione che $k \ll n$, ovvero $k = O(n)$; per cui, se la complessità dello sforzo dell'algoritmo è $\Theta(n + k)$, eseguendone la somma prevale il maggiore (e quindi lo sforzo più esoso) e quindi la **complessità diviene lineare**, ovvero $\Theta(n)$, sotto le opportune **condizioni di partenza**, ovviamente.

5.2 Bucket-Sort

L'**ordinamento per contenitori** è estremamente diffuso. Esso si basa su due **strutture dati fondamentali**: *array* e *lista sequenziale*.

Gli array, come è noto, presentano una struttura sequenziale di **celle di memoria consecutive**, avente lunghezza n . Se si considera un array A di $\text{length}(A) = n$ (con $n \geq 1$), allora il costo di posizionarsi in una qualsiasi posizione $0 \leq i \leq n$ è semplicemente $\Theta(1)$, effettuando un **accesso libero (o lineare o random)**.

Tale aspetto è estremamente significativo e vantaggioso; tuttavia, ciò comporta di sapere a priori la lunghezza dell'array (giacché si devono disporre n celle di memoria consecutive), mostrandone la natura profondamente rigida: molto spesso, però, non è nota a priori la lunghezza dei dati da memorizzare ed è più conveniente che essa cresca progressivamente nel corso dell'esecuzione dinamica dell'algoritmo: tale esigenza viene soddisfatta dalla **lista sequenziale**.

La dichiarazione di una lista L necessita di 3 fondamentali informazioni:

$$\boxed{\text{NIL} \mid \text{NIL} \mid 0}$$

in cui nelle prime due celle sono presenti due valori nulli per specificare il **puntatore al nodo iniziale** e il **puntatore al nodo finale** della lista, mentre nell'ultima cella è contenuta la **lunghezza della lista**.

Quando viene eseguita l'istruzione $\text{INSERT}(x, L)$, viene creato un nodo avente due celle: un valore (o **chiave**) e un puntatore al nodo successivo. È chiaro che il puntatore al nodo successivo della lista punta a NIL, essendo l'ultimo nodo (ossia l'unico inserito). Le due celle della lista, invece, dopo aver inserito il primo nodo, punteranno allo stesso nodo, essendo l'unico della lista, mentre il contatore viene incrementato di 1, e così via: ogni qualvolta verrà inserito un nuovo nodo, il nodo precedente punterà a quello successivo e il nuovo nodo punterà a NIL, mentre verrà aggiornato il puntatore della lista al nodo finale, così come il contatore del numero di nodi della lista.

30 Marzo 2022

Le due strutture dati fondamentali impiegabili in algoritmica sono l'array e la lista sequenziale. Si tratta in ambedue i casi di liste, ma nel caso dell'array, è fondamentale sapere a priori di quante posizioni successive si necessita per la memorizzazione. L'array viene identificato con un nome simbolico che punta alla prima cella di memoria disponibile della sequenza di n celle di memoria consecutive. Tuttavia, il costo di posizionarsi su una specifica cella di memoria per un array è $\Theta(1)$ (ovvero l'accesso ha un costo compreso tra due costanti positive $c_1 \leq c_2$): si parla, in questo caso, di **random-access** (anche definito **accesso libero o diretto**).

L'altra struttura dati prende il nome di **lista sequenziale**, per la quale non è necessario specificare a priori la lunghezza della lista, il che si rivela particolarmente utile quando non si conosce a priori la quantità di dati da dover memorizzare. La lista viene così dichiarata:

$$L \boxed{\text{NIL} \mid \text{NIL} \mid \text{O}}$$

in cui nelle prime due celle di memoria vengono memorizzati, rispettivamente, un puntatore che punta all'inizio della lista e un puntatore che punta alla fine della lista, mentre nella terza cella è presente la lunghezza della lista.

Per l'inserimento di un nuovo nodo è sufficiente impiegare la procedura $\text{INSERT}(x, L)$, grazie alla quale viene creata una doppia cella di memoria (una doppia-posizione), contenente il valore da memorizzare e un puntatore alla cella successiva (che per un inserimento in coda contiene naturalmente NIL, all'inizio). I puntatori dell'inizio e fine lista devono essere opportunamente aggiornati (entrambi puntano allo stesso unico nodo, un quanto sia nodo iniziale che finale della lista), così come il contatore della dimensione (che viene incrementato di 1), e così per ogni nuovo nodo che viene inserito.

Naturalmente, quando si necessita di cancellare un nodo in testa, si deve impiegare l'istruzione $\text{DELETE}(L)$, che provvede a cancellare il nodo in testa, aggiornando il puntatore all'inizio della lista e decrementando il contatore della lunghezza della lista: ovviamente tale operazione palesa il problema dell'**underflow** che si potrebbe verificare quando si tenta di cancellare il nodo in testa di una lista vuota, avente lunghezza nulla (vero è che come si può verificare l'underflow si può verificare l'**overflow**, anche se quest'ultimo costituisce un'eventualità assai più remota della prima).

Ciò che penalizza, tuttavia, la lista, è il fatto che il costo di accesso ad un'ipotetica cella di memoria è $O(n)$, in quanto l'**accesso può essere solamente sequenziale**.

Tuttavia, nonostante tale complessità, l'operazione di **join** favorisce la lista sequenziale anziché l'array: infatti, creare un solo array A partendo da due array A_1 e A_2 (ossia effettuando il join $A = A_1 \circ A_2$) presenta un costo $\Theta(n_1 + n_2)$, dove n_1 e n_2 sono gli elementi contenuti nei due array, in quanto bisogna inserire $n_1 + n_2$ valori all'interno di una serie di celle sequenziali. Invece, la lista sequenziale presenta un costo per il join (ossia $L = L_1 \circ L_2$) che è semplicemente $\Theta(1)$, in quanto è sufficiente aggiornare il puntatore all'ultimo nodo della lista e per la lunghezza è sufficiente eseguire la somma tra i due contatori, quindi molto più conveniente.

La lista sequenziale così descritta, che prevede un inserimento in **coda**, presenta una logica di tipo **FIFO** (first-in first-out): rassomiglia, quindi, ad una coda che presenta i metodi **ENQUEUE** per l'inserimento e **DEQUEUE** per la cancellazione. Invece, se l'inserimento è in testa, si ottiene uno **stack** (o pila), che presenta una dinamica **LIFO** (last-in first-out).

Le liste bidirezionali, invece, prevedono di impiegare due puntatori per ogni nodo (uno che punta al nodo precedente e uno che punta al nodo successivo), che quindi consentono di percorrerle la lista in due direzioni, per cui il costo di accesso diviene $O(\frac{n}{2})$.

Osservazione: La speranza matematica è sempre additiva, mentre la varianza no: una condizione sufficiente perché lo sia è l'indipendenza di eventi.

Per esempio, se si considera la funzione deterministica $f(\mathcal{X}) = \alpha \cdot \mathcal{X} + \beta$, allora si ha che

$$E[\alpha \cdot \mathcal{X} + \beta] = \alpha \cdot E[\mathcal{X}] + \beta$$

e

$$\text{var}[\alpha \cdot \mathcal{X} + \beta] = \alpha^2 \cdot \text{var}[\mathcal{X}]$$

volendo recuperare la dimensione di \mathcal{X} , è sufficiente passare allo scarto quadratico medio, o alla deviazione standard:

$$\text{sqm}[\alpha \cdot \mathcal{X} + \beta] = \sqrt{\alpha^2 \cdot \text{var}[\mathcal{X}]} = |\alpha| \cdot \text{sqm}[\mathcal{X}]$$

Se si considera un indicatore di successo I (ossia un bit aleatorio che vale 1 con probabilità p e vale 0 con probabilità $q = 1 - p$), è facile capire come $E[I] = p$, mentre $\text{var}[I] = pq$.

Se, invece, si considera, un numero n di esperimenti (ripetuti nelle stesse condizioni), a cui vengono associati n indicatori di successo $N = I_1 + I_2 + \dots + I_n$, sommando il numero aleatorio di successi N , si giunge alla distribuzione binomiale seguente, per la quale, essendo la speranza matematica additiva, così come la varianza essendo esperimenti indipendenti l'uno dall'altro:

$$E[N] = np \quad \text{e} \quad \text{var}[N] = npq$$

Volendo determinare la speranza matematica di N^2 , si può facilmente osservare che:

$$\text{var}[N] = E[(N - \mu)^2] = E[N^2] - [E[N]]^2$$

dove $\mu = E[N]$. Per cui è facile capire come

$$E[N^2] = \text{var}[N] + [E[N]]^2 \longrightarrow npq + n^2p^2 = np \cdot (q + np)$$

ma essendo la probabilità di insuccesso $q = 1 - p$ si ottiene che

$$npq + n^2p^2 = np \cdot (q + np) = np \cdot (1 - p + np)$$

Ma siccome in bucket-sort è fondamentale valutare la speranza matematica di N^2 quando la probabilità di successo $p = \frac{1}{n}$, si può facilmente capire come si semplificano la speranza matematica e la varianza, da cui

$$E[N] = np = n \cdot \frac{1}{n} = 1$$

da cui

$$np \cdot (1 - p + np) = n \cdot \frac{1}{n} \cdot \left(1 - \frac{1}{n} + n \cdot \frac{1}{n}\right) = 2 - \frac{1}{n} < 2$$

ottenenendo

$$E[N^2] = 2 - \frac{1}{n} < 2$$

Osservazione: In bucket-sort, dovendo ordinare valori reali in un ipotetico intervallo $[\alpha, \beta]$ con $\alpha < \beta$, si può facilmente considerare, per semplicità, l'intervallo $[0, 1]$, giacché la trasformazione lineare seguente

$$x_i = \frac{y_i - \alpha}{\beta - \alpha}$$

permette di ricondurre l'intervallo $[\alpha, \beta]$ all'intervallo $[0, 1]$, con $x_i \in [0, 1]$ quando $y_i \in [\alpha, \beta]$ (infatti, per $y_i = \alpha$ e $y_i = \beta$ si ottengono gli estremi dell'intervallo unitario): pertanto i due intervalli sono perfettamente equivalenti.

Tuttavia, il calcolatore non può considerare un intervallo continuo $[0, 1]$, in quanto la precisione di rappresentazione è fortemente limitata, per cui il numero di valori rappresentabili nell'intervallo $[0, 1]$ è **finito**.

Osservazione: Si osservi che i valori numerici non hanno un'unica rappresentazione, in quanto il valore 1 può anche essere scritto come $0.\bar{9}$, così come il valore 0.323 può essere scritto come 0.3229; ciò, naturalmente, quando il numero di cifre decimali a disposizione è infinito.

Osservazione: Naturalmente, l'intervallo $[0, 1]$ presenta la cardinalità del continuo, mentre il calcolatore non solo non è capace di avere una precisione infinita di rappresentazione, ma non è in grado nemmeno di garantire la precisione infinita razionale (e non continua) dell'intervallo $[0, 1]$. Ciò che è possibile fare è fissare un numero massimo di cifre decimali (come 8 nel caso di un byte), come 0.03211448.

Pertanto, considerando la legge di distribuzione di probabilità continua sull'intervallo $[0, 1]$, la probabilità di un sottointervallo $[a, b] \subset [0, 1]$ è naturalmente calcolabile come l'integrale sull'intervallo $[a, b]$ della funzione costante 1; tuttavia, tale risultato non può essere ottenuto da un calcolatore con una precisione finita, ma può essere fornita un'approssimazione di tale concetto estraendo in maniera uniforme 10^8 valori (ossia sempre un numero finito, da 0.00000001 a 0.99999999) nell'intervallo $[0, 1]$, escluso il valore estremo di destra, quindi $[0, 1[$, ciascuno dei quali presenta una

probabilità di 1 su 10^8 di essere estratto.

È intuitivo capire che lavorando con $n \gg 1$ valori da ordinare, l'algoritmo **bucket-sort** funziona particolarmente bene quando i valori da ordinare presentano una distribuzione quanto più uniforme nell'intervallo $[0, 1[$ (dimodoché non vi siano sottointervalli di omogeneità eccessivamente difforme). Di seguito si espone il listato dell'algoritmo **bucket-sort**:

Algorithm 26 Bucket-Sort

```

1: BUCKET-SORT( $A$ )
2:  $n = \text{length}(A)$ 
3: for  $i = 1$  to  $n$ 
4:   do insert  $A[i]$  in  $B[\lfloor n \cdot A[i] \rfloor]$ 
5: for  $i = 0$  to  $n - 1$ 
6:   do order  $B[i]$  with insertion-sort
7: join lists  $B[0], B[1], \dots, B[n - 1]$  in this order
```

Per l'ordinamento, quindi, oltre all'array iniziale A , si impiega un secondo array B nel quale, da 0 a $n - 1$, si inserisce inizialmente il valore NIL, e progressivamente viene creata una lista concatenata per ciascuna (non necessariamente) posizione di B (ovvero nelle posizioni dell'array B si inserirà un puntatore ad una lista sequenziale), tramite il primo ciclo for, in cui vi sono n iterazioni, per uno sforzo complessivo di $\Theta(n)$, come mostrato di seguito:

A	B
1 0.78	0 $B[0]$ → 0.17
2 0.17	1 $B[1]$ → 0.39 → 0.26
3 0.39	2 NIL
4 0.26	3 $B[3]$ → 0.78
5 0.94	4 $B[4]$ → 0.94

Pertanto, se i valori di A sono disposti uniformemente nell'intervallo $[0, 1]$, allora ogni bucket (ossia ogni posizione di B) presenterà un numero pressoché identico di valori (e siccome i bucket sono n e i numeri sono n , su ogni bucket vi saranno 0, 1 o 2 elementi). Per tale ragione è possibile impiegare l'algoritmo **insertion-sort** per ordinare i valori di ciascun bucket (anche se la complessità di insertion-sort è quadratica, per cui il prezzo da pagare è almeno $\Omega(n)$).

Dopodiché, avendo ordinato ciascun bucket, sicché i valori di ciascun bucket sono inferiori dei valori dei bucket successivi, sarà sufficiente effettuare il **join** di tutte le liste di ciascun bucket: siccome si devono effettuare $n - 1$ join con complessità $\Theta(1)$, il costo di tale operazione è $\Theta(n)$.

Osservazione: Naturalmente, l'impiego di bucket-sort è stato subordinato ad una condizione vagamente posta in principio: la supposizioni che i valori da ordinare siano disposti in maniera uniforme all'interno dell'intervallo $[0, 1]$.

Al fine di stimare la complessità di bucket-sort, si può valutare la probabilità che i valori che devono essere ordinati siano uniformemente distribuiti nell'intervallo $[0, 1[$ (in cui l'algoritmo bucket-sort funziona in maniera ottimale): in altre parole, impiegando la distribuzione uniforme, si generano a caso n numeri sull'intervallo $[0, 1[$, provvedendo ad ordinarli tramite bucket-sort, la cui complessità d'esecuzione dipende da quali numeri sono stati estratti, in altre parole si ha una **complessità aleatoria**, di cui si provvede a calcolare la speranza matematica, al fine di disporre di una stima della complessità dell'algoritmo.

Si impieghi, allora, la procedura RANDOM(0,0.99999999) per estrarre n volte, mediante la distribuzione uniforme, un numero reale nell'intervallo $[0, 1[$ e si calcoli la probabilità che tale valore si trovi inserito all'interno del bucket i -esimo, avendo n bucket (ossia n valori da ordinare); si osservi che i bucket che vengono costruiti presentano tutti la medesima ampiezza all'interno della distribuzione uniforme, per cui ogni numero presenta la stessa probabilità di finire nel bucket i -esimo,

ossia pari a

$$p = \frac{1}{n}$$

essendo n i bucket. Allora considerando N_i il numero di successi, ossia il numero di numeri che vengono inseriti all'interno del bucket i -esimo con probabilità $p = \frac{1}{n}$, per quando visto all'inizio, si ottiene che

$$E[N_i] = 1 \quad \text{e} \quad E[N_i^2] = 2 - \frac{1}{n} < 2$$

che porta a concludere che, tipicamente, in ogni bucket vi sarà contenuto uno e un solo valore (ovviamente tale risultato è aleatorio, in quanto vi è comunque una bassa probabilità che in un solo bucket vi cadano n valori, ossia la totalità, per cui la complessità pessima di insertion-sort su n elementi, verrebbe riflessa su bucket-sort).

Tralasciando la dimostrazione che $E[\Theta(f(n))] = \Theta(E())$, si calcoli la speranza matematica di una complessità, sapendo che la speranza matematica della somma è la somma delle speranze matematiche e che la complessità T dipende da tanto da n quanto da una n -upla aleatoria \mathcal{X} corrispondente alle n estrazioni di numeri casuali che sono state effettuate, per cui:

$$E[T(n, \mathcal{X})] = \Theta(n) + \sum_{i=0}^{n-1} E[T_i(\mathcal{X})]$$

in cui figura $\Theta(n)$ dal momento che il primo ciclo for presenta una complessità lineare (essendo n iterazioni); inoltre, è necessario eseguire la somma delle speranze matematiche del costo di ordinamento del bucket i -esimo tramite insertion-sort, che è naturalmente un costo aleatorio che dipende da \mathcal{X} , ossia da quali numeri sono stati estratti e quali di tali numeri sono capitati nel bucket i -esimo. Dal momento che la complessità di insertion-sort, nel caso peggiore, è quadratica, è possibile effettuare una maggiorazione per la quale, al posto della complessità $T_i(\mathcal{X})$ si inserisce $\gamma_i \cdot N_i^2$, ove γ_i è la costante che definisce la classe $\Theta(N_i^2)$ (in cui figura N_i^2 dal momento che insertion-sort viene applicato al bucket i -esimo, supponendo che la complessità sia la peggiore, quindi quadratica), ovvero:

$$\Theta(n) + \sum_{i=0}^{n-1} E[T_i(\mathcal{X})] \leq \Theta(n) + \sum_{i=0}^{n-1} E(\gamma_i \cdot N_i^2) = \Theta(n) + \sum_{i=0}^{n-1} \gamma \cdot E(N_i^2)$$

avendo sostituito γ_i con il massimo tra tutti i γ_i , ossia γ . Per cui, nel caso peggiore, siccome la speranza matematica di \mathcal{X}^2 è minore di 2, è facile capire come

$$\sum_{i=0}^{n-1} \gamma \cdot E(N_i^2) \leq n \cdot (\gamma \cdot 2)$$

avendo eseguito una somma che coinvolge n addendi ciascuno $\leq \gamma \cdot 2$. Attraverso tale approccio si può ottenere la seguente maggiorazione della speranza matematica della complessità:

$$\Theta(n) + \gamma \cdot E(N_i^2) \leq \Theta(n) + n \cdot (\gamma \cdot 2)$$

Ma ma essendo anche $n \cdot (\gamma \cdot 2)$ una funzione lineare in n , si può concludere che la **speranza matematica della complessità si bucket-sort è lineare**, da cui:

$$E[T(n, \mathcal{X})] = \Theta(n)$$

che è un risultato straordinario, pur non avendo affermato che la complessità sia lineare.

Osservazione: Si osservi che, siccome alla fine di bucket-sort si è provveduto a concatenare tutte le liste sequenziali puntate dalle posizioni dell'array B , ottenendone una lista lunga n , eseguendo uno sforzo $\Theta(n)$ è anche possibile trasformare tale lista in un array di lunghezza n con tutti i valori ordinati.

In realtà, se l'obiettivo finale è quello di ottenere un array, non serve nemmeno concatenare fra di loro tutte le liste: avendo la loro lunghezza (tramite i contatori), è sufficiente provvedere direttamente a inserire ciascuna sottolista all'interno dell'array, riservando un numero di celle consecutive già nota.

5.3 Riassunto degli algoritmi di ordinamento

Al di là degli algoritmi di ordinamento bubble-sort e insertion-sort, un primo importante algoritmo di ordinamento è merge-sort, il quale presenta una complessità nel caso peggiore log-lineare, e che prevede di adottare un approccio *dividi et impera*, in cui, tuttavia, le operazioni che vengono eseguite sono in serie, e non in parallelo. La complessità, di fatto, di merge sort, ad un passo generico n è

$$T(n) = 2 \cdot T\left(\frac{n}{2}\right) + f(n)$$

in cui la funzione $f(n)$ non è nota a priori, ma è nota solamente la sua classe Θ .

Applicando il teorema maestro, si ha che

$$T(n) = 2 \cdot T\left(\frac{n}{2}\right) + n$$

e dividendo ambo i membri per n si ottiene che

$$\frac{T(n)}{n} = \frac{T\left(\frac{n}{2}\right)}{\frac{n}{2}} + 1$$

e nel caso successivo (siccome ogni volta si divide l'array da ordinare in due parti) si ottiene che

$$T\left(\frac{n}{2}\right) = 2 \cdot T\left(\frac{n}{4}\right) + \frac{n}{2}$$

e dividendo per $\frac{n}{2}$ si ottiene:

$$\frac{T\left(\frac{n}{2}\right)}{\frac{n}{2}} = \frac{T\left(\frac{n}{4}\right)}{\frac{n}{4}} + 1$$

e così via, fino ad arrivare a

$$\frac{T(2)}{2} = T(1) + 1$$

in cui progressivamente si esprime il termine a destra dell'uguaglianza precedente con il termine successivo; infatti, nella prima uguaglianza si ha

$$\frac{T(n)}{n} = \frac{T\left(\frac{n}{2}\right)}{\frac{n}{2}} + 1$$

ma $\frac{T\left(\frac{n}{2}\right)}{\frac{n}{2}}$ è noto essere

$$\frac{T\left(\frac{n}{2}\right)}{\frac{n}{2}} = \frac{T\left(\frac{n}{4}\right)}{\frac{n}{4}} + 1$$

e così via. Per cui, dividendo n volte per 2 si ottiene che

$$\frac{T(n)}{n} = T(1) + \log_2(n)$$

in cui $\log_2(n)$ deriva dalla somma di tutti gli 1 che si ottengono sostituendo ad ogni espressione l'espressione successiva. Moltiplicando ambo i mmembri per n si ottiene che

$$T(n) = c \cdot n + n \cdot \log(n) = n \cdot \log(n)$$

ottenendo una complessità loglineare, essendo prevalente la complessità peggiore nella somma.

Nel caso di **heap-sort**, la maggior parte delle operazioni viene svolta da **heapify**, il quale verifica se viene rispettato il vincolo di maximum-heap, andando a controllare tutti i sottoalberi di ogni nodo.

Avendo costruito una catasta con il valore massimo in cima, si procede a sostituirlo con l'ultimo,

accorciando la catasta e applicando nuovamente heapify per far rispettare il vincolo, fino ad ottenere un ordinamento sui nodi dell'albero, che viene riflesso sull'array di partenza.

La complessità di build-heap è $O(n \cdot \log(n))$, così maggiorata nell'ipotesi che heapify abbia sempre complessità log-lineare. Tuttavia, tale maggiorazione potrebbe essere migliorata, giungendo ad affermare che la complessità è $\Theta(n)$, in quanto in build-heap gli heapify vengono richiamati a partire da vertici prossimi alle foglie, anzi quasi tutti i vertici sono prossimi alle foglie, per cui i sottoalberi che dovranno essere controllati saranno molto più contenuti rispetto alla totalità dell'albero di partenza: ha rilevanza specificare nel dettaglio la complessità di build-heap in quanto esso non viene impiegato solamente in heap-sort.

31 Marzo 2022

La maggiorazione della complessità di **build-heap** a $O(n \cdot \log(n))$ è esagerata, e sarebbe molto più utile dimostrare che la sua **complessità** è perfettamente **lineare**: naturalmente ciò non migliora la complessità di heap-sort che, nel suo complesso, presenta una complessità lineare; tuttavia, build-heap viene utilizzato anche in altri algoritmi, in cui la complessità lineare della procedura è rilevante nei termini della complessità globale.

Per farlo, bisogna supporre che nel caso di interesse, la derivata di una somma infinita (al più numerabile) di termini sia la somma delle derivate (che non sempre è vero nel caso di un numero di addendi infinito non numerabile). In particolare si ha che:

$$\sum_{i=0}^n q^i = \frac{1 - q^{n+1}}{1 - q} \quad \text{con } q \neq 1$$

Se $0 < q < 1$, quando $n \rightarrow +\infty$, tale successione tende ad un valore ben preciso, ossia a

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^n q^i = \lim_{n \rightarrow +\infty} \frac{1 - q^{n+1}}{1 - q} = \frac{1}{1 - q}$$

e se $q = \frac{1}{2}$ allora tale limite è 2.

Allora, derivando tale sommatoria rispetto a q (essendo una somma finita, la derivata della somma è la somma delle derivate), si ottiene

$$\frac{1}{q} \sum_{i=1}^n i \cdot q^i = \frac{(1 - q) \cdot (-(n + 1) \cdot q^n) + (1 - q^{n+1})}{(1 - q)^2}$$

E tale derivata, quando $n \rightarrow +\infty$ si traduce semplicemente in

$$\frac{1}{q} \sum_{i=1}^{\infty} i \cdot q^i = \lim_{n \rightarrow +\infty} \frac{(1 - q) \cdot (-(n + 1) \cdot q^n) + (1 - q^{n+1})}{(1 - q)^2} \rightarrow \frac{1}{q} \sum_{i=1}^{\infty} i \cdot q^i = \frac{1}{(1 - q)^2}$$

e moltiplicando ambo i membri per q si ottiene

$$\sum_{i=1}^{\infty} i \cdot q^i = \frac{q}{(1 - q)^2}$$

per cui, ancora una volta, se $q = \frac{1}{2}$ allora tale limite è 2.

Osservazione: Nell'algoritmo build-heap si deve considerare un albero binario (non necessariamente completo) sul quale viene eseguito heapify per ogni nodo (per ogni sottoalbero, vi sono due iterazioni da compiere: una per verificare se si deve operare, eventualmente fermandosi a questa se non si deve agire, l'altra per iniziare ad operare).

Al fine di ottenere una limitazione superiore del numero di iterazioni eseguite all'interno di build-heap, si suppone che l'albero binario sia completo (ovvero vi siano tutti i nodi in ogni livello), avente altezza totale H , e per ogni nodo è stato applicato heapify (anche alle foglie, per le quali non servirebbe eseguire heapify), da cui

$$\#\text{iterazioni} \leq \sum_{h=0}^H 2^{H-h} \cdot (h + 1)$$

in cui 2^{H-h} sono il numero di foglie presenti al livello h dell'albero, mentre $h + 1$ sono le iterazioni che heapify compie per ogni foglia del livello h (dovendo controllare i sottoalberi sottostanti).

Naturalmente tale sommatoria può essere riscritta come

$$\sum_{h=0}^H 2^{H-h} \cdot (h + 1) = \sum_{h=0}^H 2^{H-h} \cdot h + \sum_{h=0}^H 2^{H-h} = 2^H \cdot \sum_{h=0}^H h \cdot \frac{1}{2^h} + 2^H \cdot \sum_{h=0}^H \frac{1}{2^h}$$

e siccome si sono ottenute due sommatorie finite, si può maggiorare ulteriormente, supponendo di aggiungere tutti gli infiniti termini per ottenere delle serie infinite, ottenendo, per $q = \frac{1}{2}$:

$$\sum_{h=1}^{\infty} h \cdot \frac{1}{2^h} = \frac{q}{(1-q)^2} = 2$$

essendo la derivata della serie esposta in precedenza. Inoltre si

$$\sum_{h=0}^{\infty} \frac{1}{2^h} = \frac{1}{1-q} = 2$$

che è la serie geometrica di ragione $q = \frac{1}{2}$. Dal momento che il in ambo i casi è 2, si ottiene la maggiorazione seguente

$$2^H \cdot \underbrace{\sum_{h=0}^H h \cdot \frac{1}{2^h}}_{=2} + 2^H \cdot \underbrace{\sum_{h=0}^H \frac{1}{2^h}}_{=2} \leq 4 \cdot 2^H$$

È noto, inoltre, che l'altezza di un albero binario appartiene alla classe $\Theta(\log_2(n))$, in cui n è il numero dei nodi, da cui si evince che, essendo H l'altezza dell'albero considerato

$$4 \cdot 2^H \cong 4 \cdot 2^{c \cdot \log_2(n)} = \Theta(n)$$

in cui c è la costante superiore che deriva dalla definizione di classe $\Theta(\log_2(n))$. Ecco che allora si evince come la **complessità di build-heap** sia esattamente **lineare**.

Osservazione: Nel caso dell'algoritmo quick-sort, come in merge-sort, si costruisce un albero che viene percorso dall'estrema destra fino in basso; tuttavia, la percorrenza dell'albero, a differenza di merge-sort, avviene in maniera irregolare; dal punto di vista operativo, invece, si considera un perno (o pivot) (scelto arbitrariamente) e due confini mobili identificati da due indici: l'indice i divide i valori minori o uguali al perno da quelli maggiori, mentre l'indice j divide i valori maggiori o uguali al perno da quelli che ancora non sono stati confrontati.

Al termine del sorting, il perno viene posizionato nella sua posizione definitiva (ossia nella posizione successiva a i), ripetendo tale procedura, in maniera ricorsiva, per ciascuna delle due parti in cui l'array da ordinare viene diviso. È chiaro che se la ripartizione avviene sempre in maniera squilibrata e non omogenea (come quando l'array è già ordinato), allora la complessità di quick-sort diventa quadratica (naturalmente nel caso peggiore).

La complessità di quick-sort, nel caso medio, invece, è log-lineare: per dimostrarlo, si sarebbe dovuto supporre che i numeri da ordinare fossero tutti diversi fra di loro e calcolare la somma delle complessità di ognuna delle $n!$ permutazioni di n valori, moltiplicata per la probabilità che tale permutazione si presenti (ossia $\frac{1}{n!}$) e diviso per il numero delle permutazioni (pari a $n!$).

Tuttavia, una strada alternativa, ma ugualmente valida, è stata quella adottata, per la quale si prevedeva, di volta in volta, di estrarre casualmente il perno, dimodochè nessuna permutazione fosse favorita rispetto alle altre, e tramite il calcolo della speranza matematica, si è ottenuto che la complessità media di quick-sort è proprio log-lineare (mentre quella empirica è ottima). Inoltre, tale algoritmo per confronti, essendo un algoritmo di ordinamento generale, presenta una complessità nel caso generale e nel caso peggiore almeno log-lineare.

Osservazione: Passando ad algoritmi di ordinamento basati su strutture dati in ingresso ben precise, si è osservato che **counting-sort** presenta una complessità lineare: naturalmente, tale algoritmo si applica quando $k << n$ (anche se si potrebbe scrivere che $k = O(n)$, pur essendo una notazione fuorviante), tale per cui vi saranno da ordinare molti numeri uguali, con molte ripetizioni.

La prima fase dell'algoritmo prevede di inizializzare un array C di k posizioni a 0, per poi effettuare n iterazioni all'interno dell'array A per contare quanti valori sono più piccoli di ciascuna posizione di C : lo sforzo compiuto è, naturalmente, dato da $\Theta(n + k)$.

Effettuando, poi, un nuovo ciclo di n iterazioni (anche se al contrario, da n a 1, è irrilevante), si procede ad eseguire l'ordinamento non in loco, ma su un nuovo array B , andando progressivamente a decrementare, nell'array C , il contatore del numero di valori ancora disponibili, minori o uguali ad ogni posizione dell'array C .

Osservazione: L'algoritmo bucket-sort permette di ordinare valori reali compresi in un intervallo $[0, 1]$ (o in qualsiasi intervallo, per quanto si è visto in precedenza); ovviamente, si suppone che i valori da ordinare siano distribuiti in maniera uniforme all'interno dell'intervallo $[0, 1]$, ovvero i sottointervalli di lunghezza δ sono pressoché identici fra di loro.

Ovviamente il fatto di considerare un intervallo continuo $[0, 1]$ è una mera approssimazione, dal momento che la precisione di rappresentazione per un calcolatore è sempre finita.

Il calcolo che è stato compiuto in precedenza, durante l'analisi dell'algoritmo, non serviva a determinare la complessità media di quick-sort; anzi, in prima analisi sono stati estratti, in maniera uniforme e casuale, dei valori arbitrari nell'intervallo prescelto e poi è stata determinata la speranza matematica della complessità, permettendo di compensare l'ipotesi iniziale di uniforme distribuzione dei valori di partenza nell'intervallo prescelto. Tale calcolo ha permesso di pervenire ad un risultato molto favorevole: all'interno di ogni bucket ci si aspetta **tipicamente** di avere 1 solo dei valori di partenza, come mostrato di seguito

<i>A</i>	<i>B</i>
1 0.78	0 $B[0]$ → 0.17
2 0.17	1 $B[1]$ → 0.39 → 0.26
3 0.39	2 NIL
4 0.26	3 $B[3]$ → 0.78
5 0.94	4 $B[4]$ → 0.94

All'interno dell'array *B* saranno contenuti dei puntatori a delle liste concatenate e sequenziali: ogni valore dell'array *A*, a seconda del proprio valore, deve essere inserito all'interno di uno specifico bucket, secondo la regola seguente:

$$B[\lfloor n \cdot A[j] \rfloor]$$

Naturalmente, ciascun bucket dovrà essere ordinato per mezzo di un algoritmo di ordinamento che funziona particolarmente bene quando i valori da ordinare sono ridotti in numero, in forza dell'uniforme distribuzione presa in considerazione (come *insertion-sort*).

Alla fine, concatenando fra di loro tutte le liste di bucket, si ottiene l'array di valori perfettamente ordinamento; naturalmente, anziché concatenare le liste, siccome le liste considerate presentano dei contatori che permettono di valutare il numero di elementi in esse contenute, sarebbe possibile anche creare un array di lunghezza pari alla somma di tutti i contatori in cui inserire i valori di ogni lista, pagando sempre un prezzo $\Theta(n)$ (con n lunghezza dell'array).

6 Algoritmi sui grafi

L'origine storica della teoria dei grafi viene fatta risalire al 1736, anno in cui il matematico svizzero Euler risolse il problema dei ponti di Königsberg.

Le prime applicazioni della teoria risalgono a Kirchoff nel 1847, in relazione allo studio delle reti elettriche e a Kerli, dieci anni dopo, nello studio di problemi di chimica organica. A partire dagli anni '40, a seguito dell'avvento del computer, la teoria ha conosciuto uno sviluppo impetuoso, sollecitato da numerose e importanti applicazioni che riguardano la ricerca operativa, computer science, reti di calcolatori e di comunicazione, statistica, cluster analysis e biologia.

6.1 Problema di Eulero

Per quanto si è detto, Euler fu il primo ad introdurre il concetto di grafo per la risoluzione del problema dei ponti di Königsberg: tale città era attraversata dal fiume Pregel, che la divideva in quattro parti: due aree principali (A e B) e due isole (C e D). Le quattro zone della città erano collegate tra loro da sette ponti, come mostrato in Figura 29. Ovviamente, si usa l'imperfetto, in quanto all'epoca di Euler la città di Königsberg apparteneva alla Prussia orientale, mentre attualmente appartiene alla Russia ed è nota con il nome di Kaliningrad e, a seguito dei bombardamenti della Seconda Guerra Mondiale, alcuni dei suoi antichi ponti non esistono più oggi.

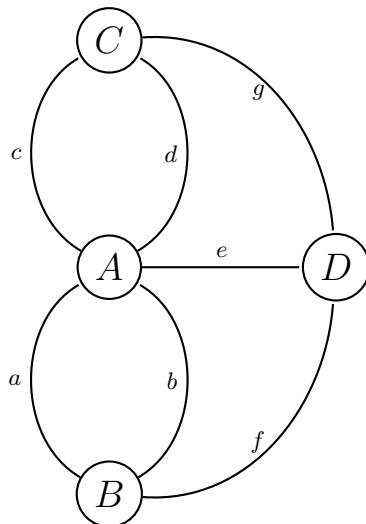


Figura 29: Problema di Eulero e dei ponti di Königsberg

Il compito che Euler si era prefisso era quello di organizzare un **percorso chiuso** che impieghi tutti i ponti una e una sola volta; dopo i suoi studi, Euler pervenne ad una soluzione molto chiara: **non esiste** alcun percorso chiuso di tale tipo.

Osservando il grafo (che non è un grafo semplice, ma un multigrafo) rappresentato in Figura 29, il collegamento tra i nodi A e C e tra i nodi A e B presenta una molteplicità doppia: in un grafo semplice non è possibile avere una molteplicità ≥ 2 .

Inoltre in un multigrafo è possibile avere dei **loop** (o *cappi*), ossia degli archi chiusi su uno stesso nodo, mentre in un grafo semplice no.

La definizione di **grado** (o **valenza**, dall'inglese *degree*) di un vertice è presto detta:

GRADO DI UN VERTICE

Il **grado** di un vertice è il numero di archi che incidono sul vertice considerato; per esempio, in Figura 29, si ha che $\deg(C) = 3$.

Mentre la definizione di **grafo connesso** è:

GRAFO CONNESSO

Un grafo, inoltre, si dice **connesso** se esiste un percorso di **archi consecutivi** che permette di collegare fra di loro ogni coppia di vertici.

E di seguito si espone la definizione di **grafo euleriano**:

GRAFO EULERIANO

Un grafo si dice **euleriano** se esiste un **percorso chiuso** (ovvero un percorso che parte e termina nel medesimo vertice) che impiega tutti gli archi una e una sola volta.

In particolare, Eulero dimostrò che un grafo è euleriano (ossia esiste un percorso euleriano) **se e solo se** è connesso e se il **grado di tutti i vertici è pari**.

Osservazione: Nell'800 **Fleury** propose un algoritmo che prevede, in prima battuta, una fase decisionale che permette di determinare se il grafo è euleriano o meno (grazie alla dimostrazione di Eulero, è facile verificare se il grafo è connesso e calcolare il grado di ogni vertice al fine di determinare se questo è pari).

Dopodichè l'algoritmo permette di determinare almeno un percorso euleriano, dal momento che ve ne potrebbero essere molteplici. La velocità d'esecuzione dell'algoritmo è significativa, in quanto sta nell'ordine del numero dei vertici e degli archi, ossia ha una complessità lineare.

6.2 Algoritmo di Hamilton

Il problema di Hamilton, invece, è il seguente: partendo da un vertice, si costruisca un percorso chiuso che unisca tutti i vertici una e una sola volta; se un percorso di tale tipo esiste, allora il grafo si dice **hamiltoniano**.

È chiaro che tale problema riguarda un grafo semplice, e non un multigrafo, in quanto avere molteplicità maggiore di 1 non è utile alla risoluzione del problema, dal momento che scegliere un arco piuttosto che un altro per collegare due nodi equivale sempre al collegamento di due nodi.

Inoltre, verificare la correttezza di una soluzione al problema è molto semplice, mentre la determinazione di una soluzione del problema è molto più complessa: si dice, infatti, che l'algoritmo è *np completo*, ovvero la risoluzione del problema non può che avvenire tramite algoritmi di complessità esponenziale (praticamente inutilizzabile); esattamente come per vertex-cover (o il problema del commesso viaggiatore), il quale è *np* completo, ci si deve accontentare di algoritmi approssimati.

6.3 Teorema delle strette di mano - Handshake theorem

Si consideri un multigrafo, tale da includere:

- molteplicità ≥ 2 per i collegamenti tra i vertici (in quanto è possibile stringere le mani più volte);
- loop (o cappi) per i vertici (in quanto una persona può stringersi le mani da solo).

In quest'ultimo caso, è fondamentale osservare come un loop, ossia un arco chiuso in uno stesso nodo, contribuisce di 2 al grado del vertice (in quanto è come se vi fossino due archi che incidono sul medesimo vertice).

Si consideri, allora, il grafo seguente:

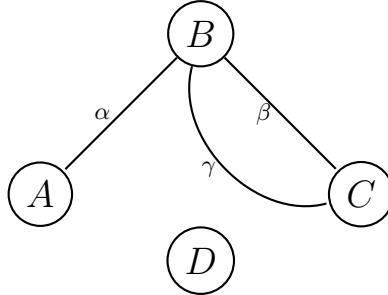


Figura 30: Problema delle strette di mano

Al fine di analizzare tale grafo, si scrivano i vertici in colonna tante volte quant'è il loro grado:

A	-	α
B	-	α
B	-	β
B	-	γ
C	-	β
C	-	γ

Analizzando la colonna di sinistra, è facile capire che essa rappresenta la somma del grado di tutti i vertici, ossia

$$\sum_{i=1}^n \deg(v_i)$$

Invece, se si considera la colonna di destra, è immediato capire come in essa ogni arco compaia due volte, in quanto presenta due estremi, ossia

$$2 \cdot |\xi|$$

Pertanto, il teorema delle afferma che **la somma del grado dei vertici è sempre pari al doppio del numero di archi**:

$$\boxed{\sum_{i=1}^n \deg(v_i) = 2 \cdot |\xi|}$$

che permette anche di capire come la somma dei gradi sia sempre un numero pari.

Osservazione: Si osservi che in un ipercubo n -dimensionale, i vertici presentano coordinate binarie di n bit (per $n = 3$ si ha 000, 001 e così via fino a 111): pertanto, è facile capire come i vertici sono in tutto 2^n (nel caso di $n = 3$, i vertici sono $2^3 = 8$).

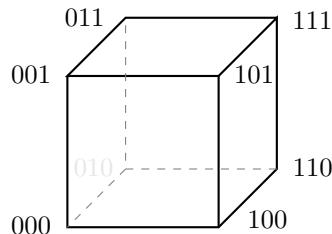


Figura 31: Ipercubo tridimensionale

Per quanto concerne gli spigoli, è fondamentale osservare come ogni spigolo collega due vertici le cui coordinate differiscono solamente per un bit (per esempio, i vertici 001 e 011 sono collegati fra di loro da uno spigolo): ciò permette di capire come, dato un vertice, è possibile creare al massimo n collegamenti da esso (ovvero vi sono n vertici ad esso adiacenti), cambiando di volta in volta una coordinata, per cui il numero totale degli spigoli è

$$\frac{2^n \cdot n}{2} = 2^{n-1} \cdot n$$

Da tale ragionamento si evince come il massimo della distanza fra due vertici è pari a n (in quanto è sufficiente alterare ad uno ad uno gli n bit di rappresentazione di ogni vertice). Se la rappresentazione di un ipercubo n -dimensionale riguardasse il collegamento tra 2^n vertici al fine di garantirne la comunicazione, con tale soluzione si avrebbe che alcune linee di comunicazione sarebbero eccessivamente occupate, favorendone delle altre.

Per la risoluzione del problema dell'eccessivo traffico su alcune linee, si potrebbe estrarre del tutto a caso un nuovo nodo il quale farà da tramite nella comunicazione tra i due nodi collegati da una linea eccessivamente trafficata: per esempio, dati A e B due nodi che comunicano frequentemente, è possibile estrarre un terzo nodo C che permetta ad A e a B di comunicare, ma indirettamente, tramite il collegamento $A - C - B$.

Tale approccio, chiaramente, potrebbe raddoppiare la distanza del collegamento di comunicazione, ma tramite un semplice calcolo di probabilità si riuscirebbe agevolmente a capire come con tale approccio il traffico si distribuirebbe in maniera uniforme su tutta la rete.

Osservazione: Il numero degli archi in un grafo semplice è sempre

$$0 \leq |\xi| \leq \frac{k \cdot (k - 1)}{2}$$

con k il numero dei vertici. Nell'ipotesi in cui il grafo dovesse essere connesso, è facile capire che se esso è completo, allora è anche connesso, per cui la limitazione di destra permane inalterata. Inoltre, si dimostrerà che un grafo connesso conterrà sempre un albero al suo interno ed è noto che un albero presenta un numero di rami pari al numero di vertici meno 1 (basta, infatti, eliminare ad uno ad uno i rami e le conseguenti foglie attaccate per accorgersene, in cui l'unica foglia che rimarrà in vita sarà la radice). Pertanto si ottiene che, in un **grafo connesso** il numero degli archi è pari a

$$\Theta(|v|) = |v| - 1 \leq |\xi| \leq \frac{|v| \cdot (|v| - 1)}{2} = \Theta(|v|^2)$$

1 Aprile 2022

Si dimostri che se due funzioni definitivamente infinite appartengono alla stessa classe di complessità Θ , allora anche i loro logaritmi appartengono alla medesima classe Θ .

In altre parole, è noto che se $f(x) = \Theta(g(x))$, allora $\exists c_1, c_2, \tilde{x} \in \mathbb{R}^+$ tali che

$$c_1 \cdot g(x) \leq f(x) \leq c_2 \cdot g(x) \quad \forall x \geq \tilde{x}$$

Allora applicando i logaritmi a tale diseguaglianza si ottiene che

$$\log(c_1) + \log(g(x)) \leq \log(f(x)) \leq \log(c_2) + \log(g(x))$$

Per cui è sufficiente considerare

$$\frac{\log(g(x))}{2} \leq \log(f(x)) \leq 2 \cdot \log(g(x))$$

essendo $g(x)$ infinita e conseguentemente anche $\log(g(x))$ è definitivamente infinito. Pertanto, $\log(g(x)) \geq \log(c_1)$ definitivamente, per qualche $x \geq \tilde{x}$.

Osservazione: Si consideri una coda, allora il metodo di inserimento di un nodo x all'interno della coda Q (inizialmente eventualmente vuota, ossia $Q = \emptyset$) prende il nome di ENQUEUE(x, Q), mentre $x = \text{DEQUEUE}(Q)$ restituisce il nodo cancellato.

Nell'utilizzo della coda, in tale corso, si farà sempre in modo che la sua lunghezza sia sempre inferiore (o eventualmente uguale) a n . Inoltre, è possibile anche impiegare un array per realizzare una **coda circolare**, in cui l'ultimo elemento punta al primo, come mostrato di seguito, nel caso di una coda di lunghezza massima 12:

$$Q = \boxed{\begin{matrix} & 1 \\ \text{NIL} & | & \text{NIL} & | & \dots & | & 15 & | & 6 & | & 9 & | & 8 & | & 4 & | & \text{NIL} \\ & 7 & & & & & & & & & & & & & & & & 11 & & 12 \end{matrix}}$$

in cui è immediato capire come la testa della coda (dall'inglese *head*) sia in corrispondenza del valore 15 (in posizione 7), mentre la coda della stessa (dall'inglese *tail*) è in corrispondenza dell'ultimo NIL (in posizione 12).

Quando la coda è vuota (ossia $Q = \emptyset$) la testa e la coda coincidono, ovvero $h = t = 1$. Al fine di accedere ai diversi elementi dell'array è necessario eseguire una **somma circolare** \oplus (ovvero la somma di due posizioni modulo la lunghezza dell'array, in questo caso $\%12$):

$$10 \oplus 4 = 14 \% 12 = 2$$

con la sola eccezione che prevede di sostituire con 12 l'eventuale resto 0 di tale operazione. Per esempio, nel caso seguente

$$Q = \boxed{\begin{matrix} & 1 \\ \text{NIL} & | & \text{NIL} & | & \dots & | & 15 & | & 6 & | & 9 & | & 8 & | & 4 & | & \text{NIL} \\ & 7 & & & & & & & & & & & & & & & & 11 & & 12 \end{matrix}}$$

l'operazione ENQUEUE(x, Q) prevede di inserire il nodo x nell'ultima posizione (ovvero 12), provvedendo ad incrementare la posizione della coda:

Algorithm 27 ENQUEUE(x, Q)

- 1: $Q[\text{tail}] = x$
 - 2: $\text{tail} = \text{tail} \oplus 1$
-

Naturalmente, nel caso preso in esame, $\text{tail} \oplus 1 = (12 + 1) \% 12 = 1$.

Se, invece, si deve cancellare un nodo in testa alla coda, si usa l'operazione DEQUEUE, la quale prevede di restituire il nodo in testa e incrementare la posizione della testa:

Algorithm 28 DEQUEUE(Q)

- 1: $x = Q[\text{head}]$
 - 2: $\text{head} = \text{head} \oplus 1$
 - 3: **return** x
-

Per conoscere la lunghezza l della coda, invece, nel caso in cui $h \leq t$, sarà semplicemente data da $l = t - h$, in caso contrario si ha $l = n - h + t$.

Naturalmente, nell'utilizzo di tale sistema, si dovrebbero controllare i due casi di *underflow* (che si ha quando si deve eseguire $\text{DEQUEUE}(Q)$ con $Q = \emptyset$ e, quindi, $l = 0$) e *overflow* (che si ha quando si deve eseguire $\text{ENQUEUE}(x, Q)$ quando $l = n$).

Tuttavia, nell'utilizzo di tale sistema, all'interno del corso, non si verificherà nessuno dei due eventi.

Osservazione: Parlando di grafi e dovendo inserire all'interno di una coda i vertici degli stessi, è facile capire come tale coda non conterrà mai, al suo interno, un numero maggiore di n nodi.

Inoltre, mentre nel caso degli algoritmi di ordinamento, la complessità di descrizione dell'input è dettata solamente da n (con n pari al numero dei valori da ordinare), ignorando completamente altri fattori di complessità (come, per esempio, la definizione del numero di cifre decimali di ogni numero da ordinare): la formulazione più corretta, in tal senso, è **complessità in funzione di n** (e non *complessità assoluta*, o *bit-complexity*).

Pertanto, nel caso dei grafi, sarebbe più conveniente provvedere a definire la complessità di inserimento dell'input **in funzione del numero dei vertici $|v|$ e del numero degli archi $|\xi|$** . A tal proposito, è bene ricordare come i grafi semplici

- non presentano loop (che contribuiscono di 2 al grado del vertice coinvolto);
- gli archi non hanno una molteplicità ≥ 2 , ma sempre unitaria o nulla;
- gli archi non presentano ordinamenti (ossia se esiste un arco $A - B$, allora esiste anche l'arco $B - A$).

Il modo più immediato al fine di descrivere l'input dei vertici e degli archi di un grafo semplice è quello di impiegare una **lista di adiacenza**. Per esempio, nel caso del grafo di seguito esposto

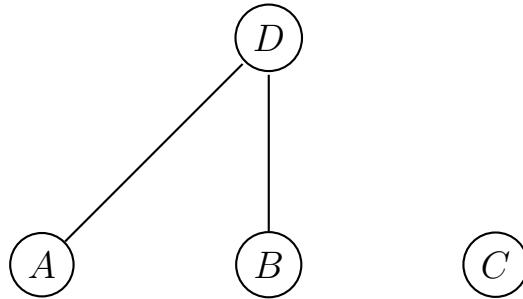


Figura 32: Descrizione di un grafo tramite una lista di adiacenza

La lista di adiacenza associata al grafo sarebbe la seguente

a	$;$	d	$;$	b	$;$	d	$;$	c	$;$	$;$	d	$;$	a	b	$;$
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

In cui, semplicemente, dato ogni nodo, si indicano tutti i nodi ad esso adiacenti (ossia collegati tramite un arco); la logica di descrizione è la seguente: si specifica il nodo interessato, si pone un separatore (“;”), si specificano tutti i nodi ad esso adiacenti e poi si pone un nuovo separatore (“;”) e dopodichè si procede con il nodo successivo.

È immediato capire come la complessità di tale descrizione dell'input (ossia il numero di caselle impiegate) è data da

$$3 \cdot |v| + 2 \cdot |\xi|$$

in quanto ogni vertice è presente 3 volte nella descrizione (una prima volta per specificare quale vertice si sta considerando, un punto e virgola ad indicare che tale vertice è stato visitato e un'ultima volta per indicare che la lista delle adiacenze di tale vertice è terminata), mentre gli archi vengono contati due volte (in quanto uno stesso arco viene considerato nella descrizione di ciascuno dei due vertici che collega); ciò porta a concludere che la complessità di descrizione dell'input, in funzione del numero di vertici e archi, è $\Theta(|v| + |\xi|)$.

Nel caso di presenza di cappi sui nodi, giacché questi incidono su un solo vertice, essi vengono contati una sola volta nella descrizione delle adiacenze, per cui si avrebbe che la complessità è

$$3 \cdot |v| + 2 \cdot |\xi| - \#\text{loop}$$

Ma si avrebbe sempre che $2 \cdot |\xi| - \#\text{loop} \geq |\xi|$, in quanto il numero dei cappi è al massimo pari al numero degli archi, per cui dal punto di vista della notazione Θ sarebbe esattamente equivalente. Nel caso gli archi presentassero molteplicità ≥ 2 , allora nella descrizione dell'input, oltre a definire il vertice interessato e la sua adiacenza, ad ogni collegamento dovrà essere associata una nuova informazione che contiene la molteplicità del collegamento, ossia il numero di archi che connettono due stessi nodi; pertanto, il numero degli archi dovrebbe essere raddoppiato, ma ancora una volta tale modifica è assolutamente ininfluente dal punto di vista della notazione Θ . In conclusione, anche combinando tutte e due le casistiche precedentemente esposte, si avrebbe sempre che la complessità della descrizione dell'input per un generico grafo, sia semplice che multigrafo, è

$$\Theta(|v| + |\xi|)$$

Osservazione: Si osservi che nella definizione della complessità di descrizione dell'input per un generico grafo, $|\xi|$ indica il numero di archi, ignorando la molteplicità, ossia il numero del tipo di collegamenti tra nodi distinti, ovviamente.

Un altro modo per descrivere l'input è la **matrice di adiacenza**, nella quale si pone 1 solamente quando si ha un arco che collega i nodi considerati, 0 altrimenti. Nel caso del grafo di Figura 32 si ha

	a	b	c	d
a	0	0	0	1
b	0	0	0	1
c	0	0	0	0
d	1	1	0	0

Tabella 2: Matrice di adiacenza

Ovviamente, nel caso di un grafo semplice, senza cappi e con molteplicità al più unitaria, tale rappresentazione è sovrabbondante: la diagonale principale è ovviamente data da tutti 0 ed è ovviamente simmetrica.

Tuttavia, se vi sono cappi, allora la diagonale principale non sarà tutta nulla e se vi è molteplicità, allora la posto di scrivere 0 e 1, si specifica il numero di archi che collegano due nodi. Naturalmente, poi, se gli archi presentano un'ordinamento, ossia sono orientati, allora semplicemente il collegamento $A - B$ verrà denotato con 1, mentre il collegamento $B - A$ verrà indicato con 0, supponendo A e B due vertici collegati da un arco che da A va in B .

Pertanto, la matrice di adiacenza è una struttura estremamente flessibile, grazie alla quale è possibile non solo descrivere gli archi, ma anche i cappi, la molteplicità e l'ordinamento; in questo caso, tuttavia, la complessità è data da $|v|^2 + 2 \cdot |v|$ (in quanto oltre alle celle si considerano anche le intestazioni di riga e di colonna), ovvero è

$$\Theta(|v|^2)$$

Tale evidenza porterebbe ad eliminare la seconda soluzione in favore della prima, in quanto, per esempio, in un albero il numero degli archi è lineare con il numero dei vertici (ovvero è $|v| - 1$), per cui $\Theta(|v| + |\xi|) = \Theta(|v|)$; se, invece, si considera un grafo denso, tale per cui il numero degli archi è

$$\frac{|v| \cdot (|v| - 1)}{2} = \Theta(|v|^2)$$

ossia è quadratico nel numero dei vertici, la complessità di descrizione dell'input è in ogni caso quadratica: pertanto, nel caso di un **grafo denso** impiegare l'una o l'altra descrizione è equivalente, mentre nel caso di un **grafo sparso**, come un albero, è più conveniente, in termini di complessità, impiegare la lista di adiacenza (anche se molto spesso si impiega la matrice di adiacenza in ogni caso, essendo esso più completa ed efficiente della lista di adiacenza).

6.4 Breadth-First-Search (BFS)

Si consideri un grafo semplice \mathcal{G} nel quale viene messo in evidenza un vertice, denominato **source** (s); per comodità di comprensione, si può interpretare \mathcal{G} come un grafo connesso, tale per cui da ogni vertice esiste un percorso che porta a qualunque altro vertice.

La visita di un grafo in ampiezza viene di seguito esposta tramite l'algoritmo **breadth-first-search**, il quale prevede di partire dal vertice sorgente s e considerare tutte le adiacenze dei vertici successivi fino a concludere la visita di tutti i vertici del grafo, scoprendo di volta in volta vertici a distanze incrementali e costruendo, di fatto, un albero di radice s che viene inserito all'interno del grafo che, se il grafo è connesso, diviene un **albero ricoprente** (dall'inglese *spanning tree*); per tale ragione la visita in ampiezza può essere impiegata per controllare se un grafo è connesso.

Al fine di semplificare la visualizzazione delle operazioni, si considera la seguente notazione:

- White: vertice mai visitato;
- Gray: vertice visitato ma non ancora abbandonato (ovvero bisogna ancora processare il vertice);
- Black: vertice visitato e abbandonato.

Di seguito si espone la fase preparatoria dell'algoritmo Breadth-First-Search:

Algorithm 29 Breadth-First-Search

```
1: BREADTH-FIRST-SEARCH( $\mathcal{G}, s$ )
2: for  $u \in v[\mathcal{G}] - \{s\}$ 
3:   do  $color[u] = \text{WHITE}$ 
4:    $d[u] = \infty$ 
5:    $\pi[u] = \text{NIL}$ 
6:    $color[s] = \text{GRAY}$ 
7:    $d[s] = 0$ 
8:    $\pi[s] = \text{NIL}$ 
9:    $Q = \emptyset$ 
```

In cui semplicemente si impostano tutti i vertici al colore WHITE (tramite l'istruzione $color[u] = \text{WHITE}$), tranne la sorgente, sulla quale si inizia ad operare; inoltre, si fa in modo che la distanza di ogni vertice dalla sorgente sia inizialmente infinita ($d[u] = \infty$) e il progenitore di ogni vertice sia NIL ($\pi[u] = \text{NIL}$), eccezion fatta per la sorgente che ha distanza nulla. Inoltre viene anche creata una coda Q nella quale verranno inseriti i vertici, ma che inizialmente è vuota, ossia $Q = \emptyset$ (che se si volesse realizzare tramite un array, si dovrebbe disporre di un array di dimensione $|v| + 1$, con $|v|$ numero dei vertici).

Taluna è la fase preparatoria, mentre di seguito si espone la continuazione del listato:

Algorithm 30 Breadth-First-Search

```
1: [...]
2: ENQUEUE( $Q, s$ )
3: while  $Q \neq \emptyset$ 
4:   do  $u = \text{DEQUEUE}(Q)$ 
5:     for  $v \in \text{Adj}[u]$ 
6:       do if  $color[v] = \text{WHITE}$ 
7:         then  $color[v] = \text{GRAY}$ 
8:            $d[v] = d[u] + 1$ 
9:            $\pi[v] = u$ 
10:          ENQUEUE( $Q, v$ )
11:         $color[u] = \text{BLACK}$ 
```

Esempio: Si consideri il grafo seguente:

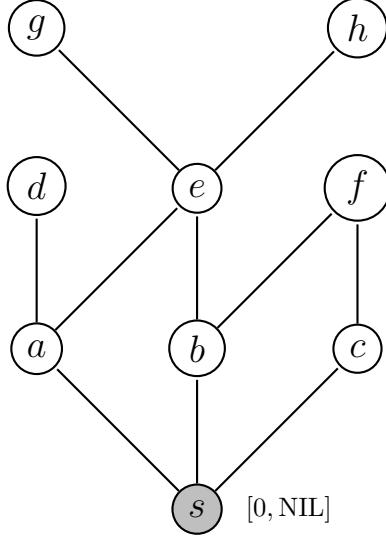


Figura 33: Esempio visita in ampiezza

Nella fase preliminare, come anticipato, tutti i vertici presentano colore WHITE, hanno distanza infinita (eccezione fatta per la radice, che ha distanza nulla e colore GRAY) e progenitore NIL; inoltre, viene creata una coda $Q = \emptyset$ nella quale viene inserito il nodo s .

Iniziando ad operare all'interno del ciclo while (fintantoché la coda non si svuota), si estrae il nodo s e si analizza tutta l'adiacenza di s , lavorando prima su a , poi su b e infine su c :

1. Siccome il vertice a è ancora WHITE, esso viene colorato di GRAY, la sua distanza dalla radice viene impostata a $d[a] = 0 + 1 = 1$ e il suo progenitore diviene $\pi[a] = s$, procedendo a mettere in coda a per una sua successiva analisi;
2. Siccome il vertice b è ancora WHITE, esso viene colorato di GRAY, la sua distanza dalla radice viene impostata a $d[b] = 0 + 1 = 1$ e il suo progenitore diviene $\pi[b] = s$, procedendo a mettere in coda anche b per una sua successiva analisi;
3. Siccome il vertice c è ancora WHITE, esso viene colorato di GRAY, la sua distanza dalla radice viene impostata a $d[c] = 0 + 1 = 1$ e il suo progenitore diviene $\pi[c] = s$, procedendo a mettere in coda anche c per una sua successiva analisi;

Avendo visitato tutta l'adiacenza di s , esso viene colorato di BLACK, ottenendo:

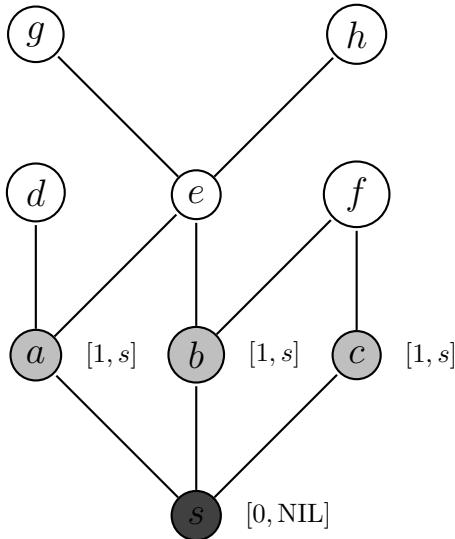


Figura 34: Esempio visita in ampiezza dopo un primo ciclo while

Procedendo, siccome la coda $Q = \{a, b, c\}$ non è vuota, si estraе il nodo a e si analizza tutta la sua adiacenza, procedendo prima con d , poi con e e infine con s :

1. Siccome il vertice d è ancora WHITE, esso viene colorato di GRAY, la sua distanza dalla radice viene impostata a $d[d] = 1 + 1 = 2$ e il suo progenitore diviene $\pi[d] = a$, procedendo a mettere in coda d per una sua successiva analisi;
2. Siccome il vertice e è ancora WHITE, esso viene colorato di GRAY, la sua distanza dalla radice viene impostata a $d[e] = 1 + 1 = 2$ e il suo progenitore diviene $\pi[e] = a$, procedendo a mettere in coda anche e per una sua successiva analisi;
3. Siccome il vertice s è BLACK, esso viene tralasciato;

Avendo visitato tutta l'adiacenza di a , esso viene colorato di BLACK, ottenendo:

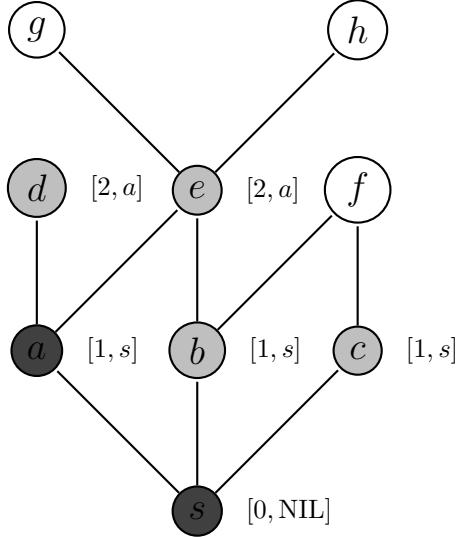


Figura 35: Esempio visita in ampiezza dopo un secondo ciclo while

Procedendo in questo modo, alla fine si ottiene il grafo seguente:

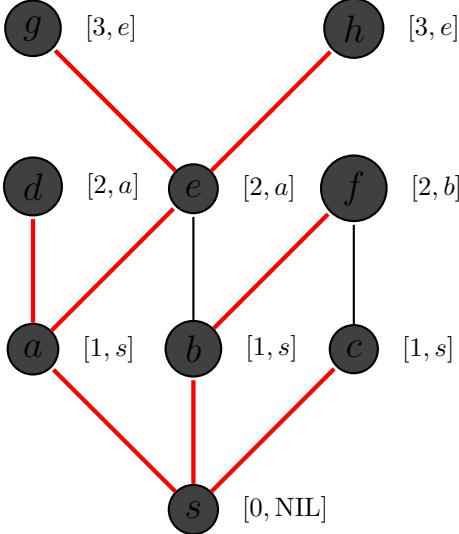


Figura 36: Esempio visita in ampiezza conclusa

Ecco che alla fine dell'esecuzione dell'algoritmo (a seguito del fatto che $Q = \emptyset$) è stato ottenuto esattamente un albero, in cui si specifica la distanza dalla sorgente e il progenitore per ciascun vertice, in cui sono evidenti i rami d'albero, evidenziati in Figura 36.

Inoltre, se tutti i vertici sono stati visitati (ovvero nessun vertice presenta ancora l'istanza ∞), è stato anche dimostrato che tale grafo è connesso; non solo, ma l'albero che ne deriva è anche il cosiddetto **spanning tree**, ovvero un albero di visita del grafo che impiega il minor numero di archi (che, nel caso di un albero, è $|v| - 1$). In altre parole, affinché un grafo sia connesso, si necessita di almeno $|v| - 1$ archi, al fine di costruire l'albero.

Osservazione: Al fine di valutare la complessità di tale algoritmo è necessario impiegare il **metodo dell'agglomerazione**, il quale permette di determinare il numero di iterazioni di cicli annidati: per esempio, se si considerano due cicli for annidati, il primo di lunghezza n e il secondo di lunghezza m , allora il numero totale di iterazioni è $n \times m$; ma se le iterazioni del secondo ciclo for sono variabili, e sono m_i (ossia dipendono da i), allora, ovviamente, le iterazioni eseguite sono $m_1 + m_2 + \dots + m_n$.

Naturalmente, giacché in tale algoritmo il ciclo while esterno esegue un'iterazione per ognuno degli n vertici, mentre il ciclo for interno esegue un numero di iterazioni pari al numero di vertici nell'adiacenza di ogni vertice, è facile capire come il numero di iterazioni totali dell'algoritmo è

$$\#\text{iterazioni} = \sum_{i=1}^n |\text{Adj}(u_i)|$$

Ma naturalmente, sommare le adiacenze di un vertice significa sommare il grado di ciascun vertice, da cui

$$\#\text{iterazioni} = \sum_{i=1}^n |\deg(u_i)|$$

Ma è anche noto che la somma dei gradi di ciascun vertice è pari a $2 \cdot |\xi|$, in quanto per ogni nodo, uno stesso arco viene contato due volte, per cui

$$\#\text{iterazioni} = 2 \cdot |\xi|$$

Nell'ipotesi in cui il grafo \mathcal{G} sia connesso, è noto che $|\xi| \geq |v| - 1$, per cui $|\xi| = \Omega(|v|)$. Per cui si ha che

$$|v| + |\xi| - 1 \leq 2 \cdot |\xi| \leq 2 \cdot (|v| + |\xi|)$$

che permette di concludere come la complessità dell'algoritmo sia lineare, ovvero $\Theta(|v| + |\xi|)$; cioè, naturalmente, nell'ipotesi in cui il grafo \mathcal{G} sia connesso: se esso non fosse connesso, invece, il numero delle operazioni da eseguire potrebbe diminuire anche drasticamente (in quanto l'algoritmo BFS opera fintantochè c'è un percorso di nodi consecutivi da analizzare), il che porta a concludere come la complessità della visita in ampiezza di un ipotetico grafo \mathcal{G} sia $O(|v| + |\xi|)$.

6.5 Depth-First-Search (DFS)

Per visualizzare la logica dell'algoritmo di visita dei vertici di un grafo in profondità, si parta dall'ipotesi che il grafo da visitare sia un albero: più precisamente, l'obiettivo è quello di allontanarsi dalla sorgente il prima possibile, andando a percorrere il grafo dalla distanza maggiore, per poi tornare indietro.

Osservazione: L'algoritmo di Fleury per la risoluzione del problema di Eulero prevede di verificare se esista un percorso euleriano per visitare tutti i nodi del grafo impiegando tutti gli archi una e una sola volta.

Naturalmente, nell'ipotesi in cui il **grafo sia connesso**, si procede a decrementare tutti i gradi dei vertici ogni volta che si esce o si entra in uno di tali vertici, fintantochè le valenze residue degli stessi non sono tutte nulle. Si capisce, dunque, per costruzione, come tale problema abbia soluzione solo quando i **gradi dei vertici sono tutti pari** (in quanto entrando e uscendo ogni volta da ogni nodo, il grado diminuisce di 2, per cui esso deve essere necessariamente pari).

In altre parole, alla base dell'algoritmo di Fleury si pone la tecnica che prevede di costruire dei mini-cicli euleriani, connettendoli fra di loro fino a che non si esaurisce l'intero grafo (se vi sono valenze ancora non nulle, essendo il grafo connesso, significa che vi sono archi ancora inesplorati).

L'algoritmo BFS che permette di effettuare la visita di un grafo in ampiezza è concettualmente molto semplice, ma allo stesso estremamente funzionale. Il listato precedentemente esposto prevede che il grafo non sia orientato (ovvero gli archi non presentano uno specifico ordinamento) e in esso deve essere messo in evidenza uno specifico vertice (ossia la *sorgente*), la quale è la radice dell'albero che si va costruendo. La logica alla base della visita in ampiezza prevede di analizzare l'adiacenza dei vertici che via via vengono scoperti a partire dalla sorgente (se il loro colore è WHITE, ossia non sono ancora stati scovati). Quando la lista dei vertici è terminata, è terminata anche la visita del grafo, che non è necessariamente stato visto nella sua interezza, in quanto altri nodi potrebbero essere inclusi in una componente连通的 del grafo stesso, ma di cui la sorgente non fa parte e, quindi, sono irraggiungibili seguendo un percorso di nodi consecutivi e adiacenti.

Un modo alternativo di esecuzione dell'algoritmo BFS è quello che prevede di non mettere in evidenza una sorgente, ma partire dal primo nodo libero che compare nella descrizione del grafo ed eseguire la sequenza di operazioni note fino ad esaurire tutta la componente连通的 (e quindi costruire l'albero di visita) di cui fa parte il vertice preso in esame. Dopodiché si considera un secondo vertice libero nella descrizione del grafo e si analizza tutta la componente连通的 del grafo ad esso associata, fino a quando non vi sono più vertici liberi nella descrizione a disposizione e, quindi, tutto il grafo, anche se non connesso, è stato visitato. In questo modo, però, vengono create più alberi di visita, uno per ciascun nodo libero considerato nella descrizione del grafo (ossia per ogni componente连通的).

FORESTA

Quando un grafo si compone di tanti alberi di visita, allora si dice che tale grafo è una **foresta** (dall'inglese *forest*).

Per cui, inevitabilmente, un grafo non connesso viene coperto da una foresta, ossia da una moltitudine di alberi di visita, uno per ciascuna componente连通的.

Inoltre, nel caso in cui il grafo da visitare sia orientato (ovvero non si hanno archi, ma frecce), l'esecuzione dell'algoritmo rimane invariata: l'unica variante riguarda l'adiacenza di un vertice v che deve essere considerata come l'insieme dei nodi che sono collegati a v tramite una freccia che da v giunge fino al vertice adiacente (ovvero **frecce uscenti** da v), e non il viceversa.

Osservazione: Si osservi che in un grafo orientato (**directed graph**), il problema della connessione del grafo è molto più delicato rispetto ad un grafo non orientato, come mostrato di seguito:

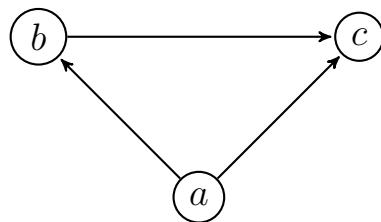


Figura 37: Esempio grafo orientato

in cui, partendo da a , si riescono a raggiungere tutti i nodi, mentre partendo da b si riesce a araggiungere solamente il nodo c .

In altri termini, in un grafo connesso non ha significato parlare di componenti connesse (in quanto dipende dal vertice da cui si parte): ciò si cui ha senso parlare sono i vertici che è possibile raggiungere partendo da un dato nodo (seguendo l'opportuna orientazione degli archi).

GRAFO ORIENTATO FORTEMENTE CONNESSO

Un **grafo orientato** si dice **fortemente connesso** se da ogni vertice è possibile raggiungere ogni altro vertice (anche se il percorso di andata potrebbe essere diverso da quello del ritorno, per cui i **percorsi non si possono, in generale, invertire**).

Inoltre, anche la differenza tra albero orientato e non orientato è molto importante: infatti, in un albero radicato, anche se inserito all'interno di un grafo non orientato, è immediato osservare come sia privilegiato il cammino che permette, dalla radice, di raggiungere tutte le foglie dello stesso.

Osservazione: Si osservi che nella descrizione di un grafo orientato, non è più detto che uno stesso collegamento tra due nodi sia ravvisabile sia nell'adiacenza dell'uno che dell'altro, essendo il grafo orientato; pertanto, nella definizione della complessità di descrizione, non comparirà più l'addendo $2 \cdot |\xi|$, ma $|\xi|$ soltanto. In ogni caso, però, la complessità di descrizione permane $\Theta(|v| + |\xi|)$.

Notazione: Si osservi che se in un dato nodo vi sono n frecce uscenti e m frecce entranti, allora si dice che l'**out-degree** è n , mentre l'**in-degree** è m .

La visita in profondità di un grafo orientato avviene analizzando di volta in volta i vertici ancora liberi nella descrizione del grafo stesso, senza mettere in evidenza una sorgente. Inoltre, è immediato capire come l'intero grafo verrà ricoperto da una foresta, ossia da un insieme di alberi costruiti raggiungendo quanti più vertici possibili seguendo l'orientazione degli archi di collegamento.

Per quanto già anticipato, la visita in profondità avviene seguendo una logica ricorsiva, cercando di allontanarsi il più possibile dal vertice preso in esame; inoltre, tale algoritmo non compie alcun tentativo al fine di determinare le distanze dei nodi dal vertice preso in esame a differenza del metodo di visita per ampiezza (infatti nell'algoritmo la d impiegata sta a denotare **discovery**, ovvero scoperta).

Com'è noto, l'algoritmo DFS opera senza specificare la sorgente del grafo sulla quale si sta lavorando, come mostrato di seguito:

Algorithm 31 Depth-First-Search

```

1: DEPTH-FIRST-SEARCH( $\mathcal{G}$ )
2: for  $u \in \mathcal{V}$ 
3:   do  $color[u] = \text{WHITE}$ 
4:    $\pi[u] = \text{NIL}$ 
5:   time = 0
6: for  $u \in \mathcal{V}$ 
7:   do if  $color[u] = \text{WHITE}$ 
8:     then DEPTH-FIRST-SEARCH-VISIT( $u$ )

```

In cui \mathcal{V} rappresenta l'elenco dei vertici nell'ordine in cui compaiono nella descrizione del grafo. Inoltre, con $time$ si fa riferimento all'orologio temporale che permette di rilevare a che istante vengono scoperti e poi analizzati i vertici del grafo.

È ovvio, inoltre, che la complessità esecutiva fino all'istruzione 5 è $\Theta(|v|)$.

Di seguito, invece, si espone il listato della procedura Depth-First-Search-Visit, ossia la visita in profondità di un generico vertice u , nella quale vi sono 3 parametri che vengono considerati:

1. l'orologio temporale $time$;
2. l'istante di scoperta di un vertice, denotato con d , ovvero *discovery*;
3. l'istante di abbandono di un vertice, denotato con f , ovvero *final*.

Algorithm 32 Depth-First-Search-Visit

```

1: DEPTH-FIRST-SEARCH-VISIT( $u$ )
2:  $color[u] = \text{GRAY}$ 
3: time = time+1
4:  $d[u] = \text{time}$ 
5: for  $v \in \text{Adj}[u]$ 
6:   do if  $color[v] = \text{WHITE}$ 
7:     then  $\pi[v] = u$ 
8:       DEPTH-FIRST-SEARCH-VISIT( $v$ )
9:    $color[u] = \text{BLACK}$ 
10:  time = time+1
11:   $f[u] = \text{time}$ 

```

In cui $d[u] = \text{time}$ è da interpretarsi come l'istante di scoperta del vertice u , ossia $d[u]$, è pari al tempo segnato dall'orologio temporale time.

Analogamente $f[u] = \text{time}$ è da interpretarsi come l'istante di abbandono del vertice u , ossia $f[u]$, che è sempre pari al tempo segnato dall'orologio temporale time.

Nella rappresentazione grafica, ogni vertice verrà corredata da due informazioni $[d, f]$, in cui d rappresenta l'istante di scoperta, mentre f quello di abbandono.

Esempio: Si consideri il grafo seguente:

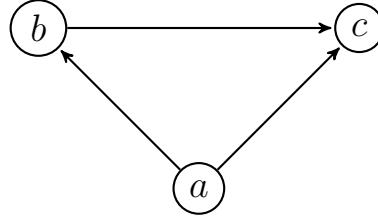


Figura 38: Esempio di visita in profondità

Allora, seguendo il listato, partendo dal vertice a , si ottengono i 3 seguenti stati, all'istante temporale time = 1, time = 2 e time = 3:

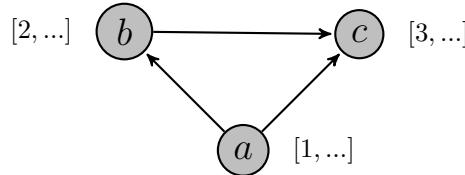
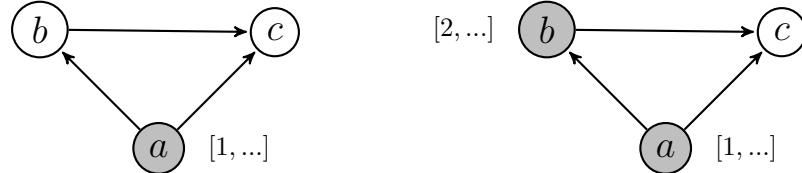


Figura 39: Esempio di visita in profondità con scoperta di tutti i vertici

E siccome l'adiacenza di ciascun vertice, ora, è completamente grigia, ognuno dei nodi visitati viene abbandonato (ripercorrendo il cammino al contrario), rispettivamente agli istanti time = 4, time = 5 e time = 6:

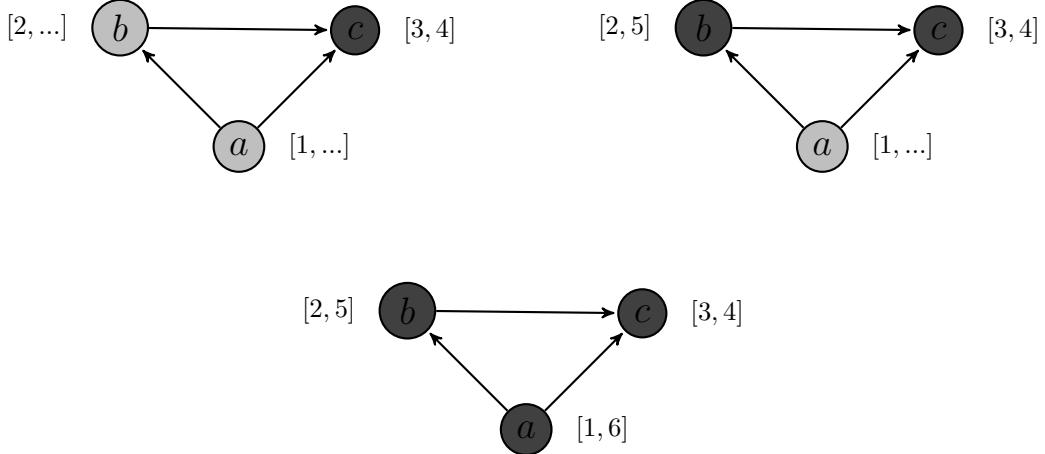


Figura 40: Esempio di visita in profondità con abbandono di tutti i vertici

Osservazione: Si osservi che, nel caso in cui nella descrizione del grafo, il primo vertice nell'elenco fosse stato b e non a , allora, alla fine della procedura di visita, si sarebbe ottenuto il grafo seguente:

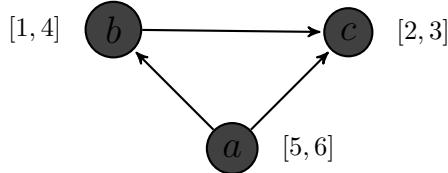


Figura 41: Esempio di visita in profondità alternativo

Ovviamente, gli istanti di visita e abbandono sono differenti rispetto all'esempio precedente; ma non solo, dal momento che partendo da b , non vi era modo di raggiungere a a causa dell'orientazione degli archi, mentre nel primo caso vi era stato bisogno di un solo albero di visita, in questo caso ne sono serviti due.

Infatti, con una visita in profondità non solo può cambiare la foresta a seconda della casualità della descrizione del grafo, ma può cambiare anche il numero degli alberi che popolano la foresta, a seconda di come vengono descritti i nodi del grafo.

come mostrato nel seguente esempio:

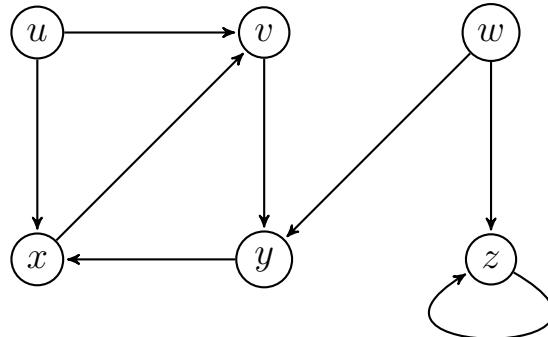


Figura 42: Esempio di visita in profondità più complesso

Nell'ipotesi in cui la descrizione dei vertici del grafo avvenga in ordine alfabetico, si parte da u e si descrive il primo albero di visita:

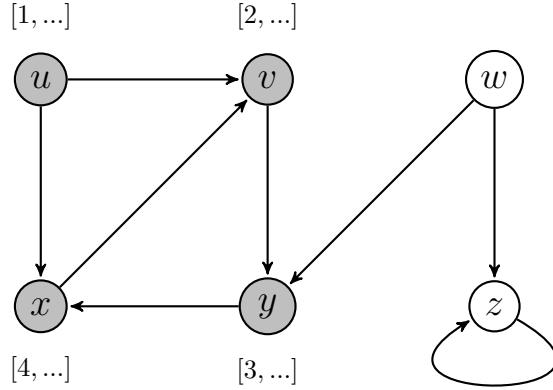


Figura 43: Esempio di visita in profondità più complesso dopo una prima scoperta

Dal momento che il nodo x presenta tutta la sua adiacenza grigia, si termina la visita e si ripercorre a ritroso il cammino compiuto (eseguendo ricorsivamente quanto già fatto, analizzato, quindi, ogni volta, l'adiacenza dei nodi su cui ci si trova), aggiornando il colore dei vertici e anche l'istante di abbandono f , mettendo in evidenza i rami d'albero:

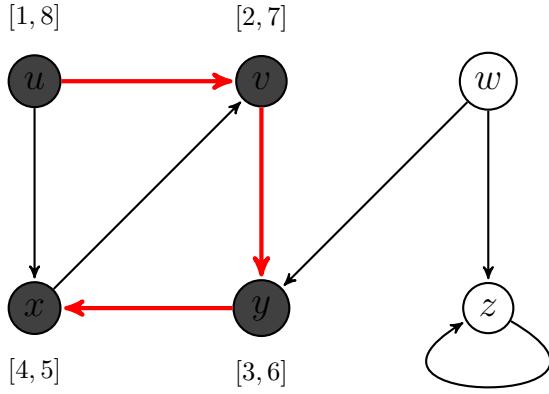


Figura 44: Esempio di visita in profondità più complesso dopo l'abbandono

A questo punto, visto che vi sono ancora vertici WHITE, è necessario considerare un nuovo albero di visita, che origina dal vertice w (che è il primo vertice libero nella descrizione del grafo) e prosegue al vertice z :

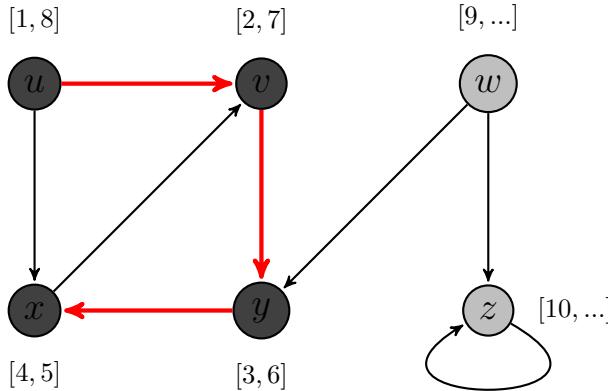


Figura 45: Esempio di visita in profondità più complesso dopo una seconda visita

Dal momento che l'adiacenza di z è grigia, si termina la visita e si ripercorre a ritroso il cammino compiuto (eseguendo ricorsivamente quanto già fatto, analizzato, quindi, ogni volta, l'adiacenza

dei nodi su cui ci si trova), aggiornando il colore dei vertici e anche l'istante di abbandono f , mettendo in evidenza i rami d'albero:

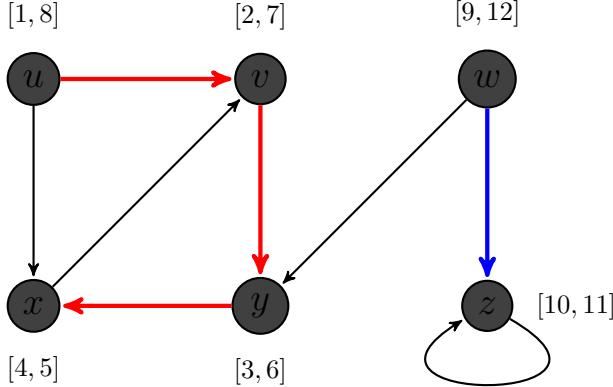


Figura 46: Esempio di visita in profondità più complesso dopo l'abbandono definitivo

Osservazione: Ecco che l'intero grafo è stato totalmente coperto da una foresta composta da due alberi.

Inoltre, è importante anche osservare che, nonostante con l'algoritmo di visita in ampiezza si conoscevano le distanze e le profondità (nonché i progenitori), all'interno di tale algoritmo sono presenti comunque i due valori (d e f) che devono essere opportunamente considerati al fine di ricostruire la foresta di vista.

6.6 Teorema delle parentesi

Il teorema delle parentesi prevede di considerare solamente le combinazioni sintatticamente corrette, ossia aventi significato dal punto di vista sintattico, delle parentesi tonde e graffe; tali combinazioni corrette sono:

$$(\square) \quad [\square] \quad \square() \quad ()\square$$

Interpretando la parentesi aperta come d (ossia l'istante di scoperta di un nodo) e la parentesi chiusa come f (ossia l'istante di abbandono di un nodo) si possono formulare le seguenti osservazioni: quando si visita il nodo u e poi il nodo v si ottiene che

$$d(u) < d(v) < f(v) < f(u)$$

e quando si visita prima v e poi u , invece, si ottiene che

$$d(v) < d(u) < f(u) < f(v)$$

In questo modo, associando le parentesi \square al nodo v e le parentesi $()$ al nodo u si ottiene

1. (\square) , che significa che il vertice quadro è discendente dal vertice tondo (nell'albero, o foresta, costruiti);
2. $[\square]$, che significa che il vertice tondo è discendente dal vertice quadro (nell'albero, o foresta, costruiti).

Naturalmente, si sarebbero potute considerare anche altre due configurazioni: $\square()$ e $()\square$, da cui non si può evincere nessuna relazione di parentela (che potrebbe comunque esserci, ma non è così identificabile).

6.7 Teorema del cammin bianco

Di seguito si espone il **teorema del cammin bianco**:

TEOREMA DEL CAMMIN BIANCO

Si supponga di considerare due vertici u e v , tale per cui $d(u) < d(v)$ (ovvero all'istante $d(u)$, il vertice v non è ancora stato scoperto): allora, se esiste un **cammino bianco** tra i due vertici (ovvero esiste un percorso composto solamente da vertici bianchi consecutivi che permette di collegare u con v), il nodo v sarà **necessariamente un vertice discendente** da u : ciò non significa che u e v verranno poi collegati, nell'albero che si andrà a costruire, tramite il cammino bianco individuato, ma significa che verrà impiegato un qualsivoglia cammino che porta da u a v , essendo v discendente di u .

Taluna è una condizione necessaria (e sufficiente), in quanto all'istante $d(u)$, partendo da u , l'unico modo per raggiungere il suo discendente v , è proprio tramite un cammino bianco (secondo l'algoritmo).

Pertanto, se v è un discendente di u , allora necessariamente all'istante di tempo $d(u)$ esiste un cammino bianco tra u e v ; tuttavia non è immediato da capire il contrario, ovvero che se esiste un cammino bianco tra u e v , allora v è un discendente di u , ed è proprio ciò che afferma il teorema.

Dimostrazione: Si supponga che, all'istante $d(u)$, il vertice v non sia stato scoperto e che esista un cammino bianco tra u e v ; si proceda **per assurdo** e si supponga che v non sia un vertice discendente di u , mentre un vertice w tale che $d(u) < d(w) < d(v)$ invece sia un suo discendente:

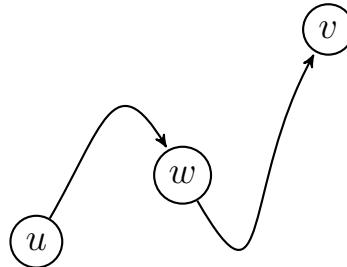


Figura 47: Dimostrazione teorema del cammin bianco

Allora, siccome w è discendente di u , applicando il teorema delle parentesi si ottiene che

$$d(u) \leq d(w) < f(w) \leq f(u)$$

in cui è necessario porre il segno di minore uguale in quanto non è possibile escludere che il primo vertice discendente di u e diverso da v sia proprio il vertice u (ossia $w = u$).

Gungendo al vertice w all'istante $d(w)$, si esamina l'adiacenza di w , ma non è determinabile a priori l'istante di tempo in cui si esaminerà l'arco $w - v$, ossia $d(v)$; in tale istante di tempo, si possono palesare tre casi possibili

- il colore di v è BLACK, per cui si ha che $f(v) < f(w)$, per cui v è un discendente di u , dal momento che dal teorema delle parentesi si ha:

$$d(u) < d(v) < f(v) < f(w) \leq f(u) \rightarrow d(u) < d(v) < f(v) < f(u)$$

dal momento che per ipotesi, all'istante $d(u)$ vi era un cammino bianco tra u e v , e quindi $d(u) < d(v)$ e per quanto supposto in questo primo caso, $f(v) < f(w) \leq f(u) \rightarrow f(v) < f(u)$.

- il colore di v è GRAY, per cui si ha che $d(v) < f(w)$, ma allora v è ancora un discendente di u , in quanto sempre dal teorema delle parentesi si ha

$$d(u) < d(v) < f(w) \leq f(u)$$

e quindi l'intervallo $[d(v), f(v)]$ è contenuto interamente all'interno dell'intervallo $[d(u), f(u)]$ (dovendo essere sintatticamente corretto), per cui si ha

$$d(u) < d(v) < f(v) < f(u)$$

dal momento che per ipotesi, all'istante $d(u)$ vi era un cammino bianco tra u e v , e quindi $d(u) < d(v)$.

3. il colore di v è WHITE, ma allora da w si passa a v e quindi è implicito come v sia un discendente di u , essendo un discendente di w che è un discendente di u ; per cui, sempre per il teorema delle parentesi:

$$d(u) \leq d(w) < d(v) < f(v) < f(w) \leq f(u)$$

In tutti i casi possibili, pertanto, si è visto come v sia in ogni caso un discendente di u , il che contrasta con l'ipotesi assurda assunta in principio.

Osservazione: Il problema noto come **topological sorting** riguarda i **grafo orientati aciclici** (dall'inglese DAG - Directed Acyclic Graph, in cui directed indica l'orientazione degli archi), ossia privo di cicli (considerando come ciclo un percorso che, partendo da un vertice, permette di ritornare al medesimo vertice di partenza, diverso da percorsi di lunghezza nulla). Esempi di ciclo viene proposto di seguito:

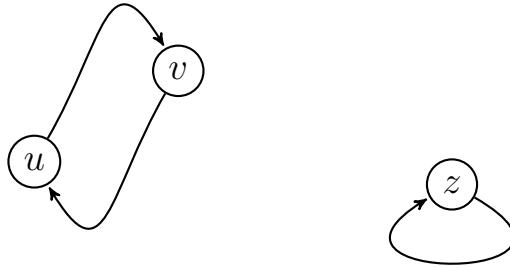


Figura 48: Esempi minimali di ciclo in un grafo orientato

Un altro nome con cui sono noti i grafi orientati aciclici (o DAG) è **grafo di propedeuticità**, ossia un grafo in cui si ha una gerarchia fra i vertici, ossia non è possibile passare da un vertice a qualsiasi altro, ma bisogna rispettare, appunto, delle **propedeuticità**, come mostrato di seguito:

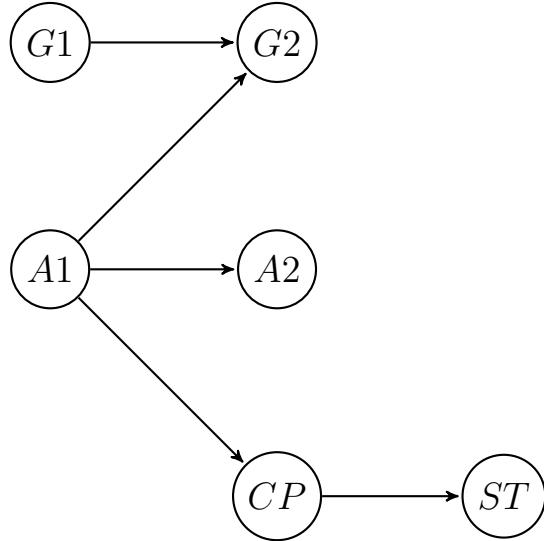


Figura 49: Esempio di grafo di propedeuticità

All'interno di tale grafo, naturalmente, non è possibile considerare dei cicli, perché altrimenti non si potrebbero rispettare le propedeuticità. Al fine di visitare la totalità di un DAG, è possibile impiegare una visita di profondità, la quale consente di rispettare le propedeuticità, come mostrato di seguito, nell'ipotesi in cui nella descrizione del grafo si parta da $G2$, il quale precede $A1$ che è seguito da $G1$, per cui si ottiene, alla fine della visita:

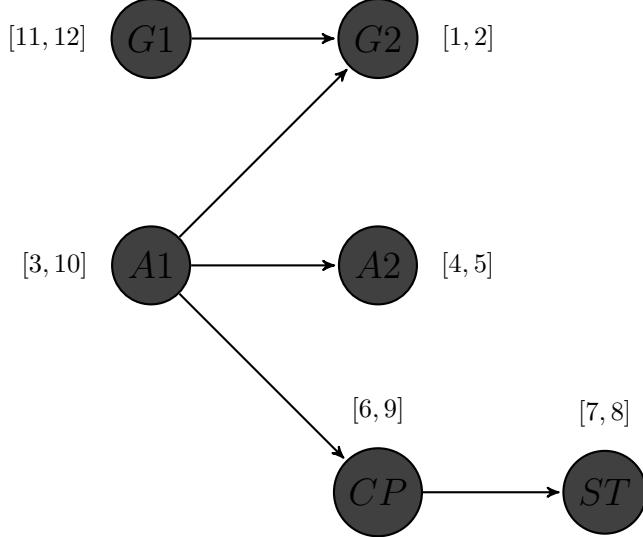


Figura 50: Esempio di visita in profondità di un grafo di propedeuticità

Avendo completato la visita in profondità, si considerino, ora, gli istanti di abbandono (f) dei vertici al contrario, ponendoli in sequenza dal maggiore al minore; si ottiene:

$$G1 - A1 - CP - ST - A2 - G2$$

Ecco che in questo modo si rispettano tutte le propedeuticità previste dal DAG esposto in Figura 49:

- $G1$ precede $G2$
- $A1$ precede $G2$
- $A1$ precede $A2$
- $A1$ precede CP
- CP precede ST

Ciò permette di capire come l'ordinamento topologico di un DAG possa essere effettuato tramite una visita in profondità.

6.8 Teorema di classificazione degli archi

Si consideri una visita in profondità e l'analisi di un arco da u in v

$$u \longrightarrow v$$

Allora, per il teorema delle parentesi, si possono avere diverse configurazioni sintatticamente corrette, che si possono o meno presentare anche concretamente:

1. La configurazione seguente

$$d(u) < f(u) < d(v) < f(v)$$

ammissibile dal teorema delle parentesi, in realtà **non è compatibile** con l'esistenza dell'arco tra u e v : prima di terminare l'analisi di u (ossia $f(u)$) è necessario analizzare tutta la adiacenza di u e quindi anche lo stesso v , per cui $d(v) < f(u)$.

2. La configurazione seguente

$$d(v) < f(v) < d(u) < f(u)$$

ammissibile dal teorema delle parentesi, è anche una configurazione lecita e compatibile con la visita in profondità: precedentemente alla scoperta di u , tramite cammini alternativi, è stato scoperto v e terminata l'analisi del sottoalbero innestato in v , per poi scoprire u e terminare l'analisi del sottoalbero innestato in u (e in tale analisi si scopre anche l'arco $u \rightarrow v$, ma essendo v già stato analizzato, esso è BLACK).

Tuttavia, da tale configurazione (che, per il teorema delle parentesi si può interpretare come $\square()$) non si può determinare la discendenza di nessuno dei due nodi u e v (i due sottoalberi “non si parlano”): ciò significa che l'arco $u \rightarrow v$ è un **arco di taglio** (dall'inglese **cut**), mostrato di seguito:

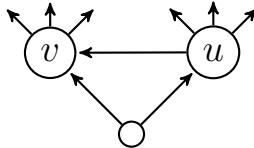


Figura 51: Esempio arco di taglio

3. La configurazione seguente

$$d(u) < d(v) < f(v) < f(u)$$

ammissibile dal teorema delle parentesi, è anche una configurazione lecita e compatibile con la visita in profondità; inoltre, da ciò si può capire come la scoperta dell'arco $u \rightarrow v$ possa avvenire in due situazioni

- (a) se $d(u) < d(v)$, allora l'arco $u \rightarrow v$ è un **arco bianco**, in quanto v non era ancora stato scoperto, quindi è WHITE: tale arco è destinato a divenire un **ramo d'albero**;
- (b) se $f(v) < f(u)$, allora l'arco $u \rightarrow v$ è un **arco nero**, in quanto dopo aver terminato l'analisi di v , analizzando u e la sua adiacenza, il vertice v sarà BLACK: taluna è la configurazione che prevede che l'arco $u \rightarrow v$ punti ad un discendente, pertanto prende il nome di **arco in avanti** (dall'inglese **forward**).

4. La configurazione seguente

$$d(v) < d(u) < f(u) < f(v)$$

ammissibile dal teorema delle parentesi, è anche una configurazione lecita e compatibile con la visita in profondità; in questo caso, in particolare, avendo già scoperto v prima di u , quando durante l'analisi dell'adiacenza di u , si considera l'arco $u \rightarrow v$, v è GRAY, per cui l'arco guarda ad un progenitore e, per tale ragione, prende il nome di **arco all'indietro** (dall'inglese **backwards**).

Osservazione: Supponendo che il grafo non sia orientato, ovvero vi è sempre, oltre che all'arco $u \rightarrow v$, anche l'arco $v \rightarrow u$, la prima come la seconda configurazione non sono ammissibili, semplicemente per simmetria (la prima non è compatibile con l'arco $u \rightarrow v$, la seconda con l'arco $v \rightarrow u$).

Nella terza configurazione, l'unica delle due situazioni che si può verificare è la (a), in quanto la (b) non è compatibile con l'arco $v \rightarrow u$. Ovviamente, nel caso di un grafo non orientato, la situazione (a) si configura come un **ramo d'albero**. L'ultima configurazione è ammissibile anche se il grafo è non orientato, in quanto simmetrica rispetto alla precedente e si configura come un **arco all'indietro** che segnala la presenza di cicli.

7 Aprile 2022

Gli archi orientati, ossia i rami degli alberi di una foresta, si pongono alla base della visita in profondità.

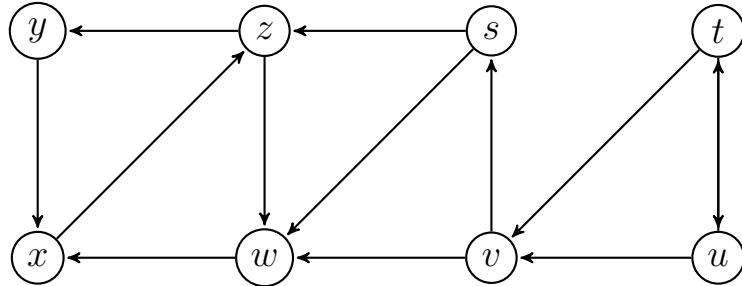


Figura 52: Esempio con visita in profondità

Per l'analisi di tale grafo si parte dal vertice che per primo compare nella descrizione del grafo. Nel caso considerato, si considera come 1° il vertice s , giacché la descrizione è $s|z$; passando a z si scopre tale vertice per 2°; conoscendo la descrizione $x|y$ si procede a scoprire y per 3°; seguendo la descrizione $y|x$ si scopre il nodo x per 4°; dal momento che il collegamento tra x e z prevede la scoperta di un nodo già noto, all'istante 5° si termina l'analisi della adiacenza di x ; tornando indietro al nodo y , si termina l'analisi della sua adiacenza all'istante 6°, ingrossando il collegamento perché è un *ramo d'albero*; allo stesso modo, si ritorna indietro al nodo s , ingrossando ancora una volta il ramo d'albero.

... continua ...

Osservazione: Analizzando su un diagramma temporale gli istanti di scoperta e di terminazione dell'analisi di ciascun verticce si ottiene:

1	2	3	4	5	6	7	8	9	10
$(s$	$(z$	$(y$	$(x$	$x)$	$y)$	$(w$	$w)$	$z)$	$s)$

Dopodiché si deve anche descrivere anche l'albero che successivamente viene descritto:

11	12	13	14	15	16
$(t$	$(v$	$v)$	$(u$	$u)$	$t)$

Osservazione: Da ciò si può descrivere un nuovo grafo;

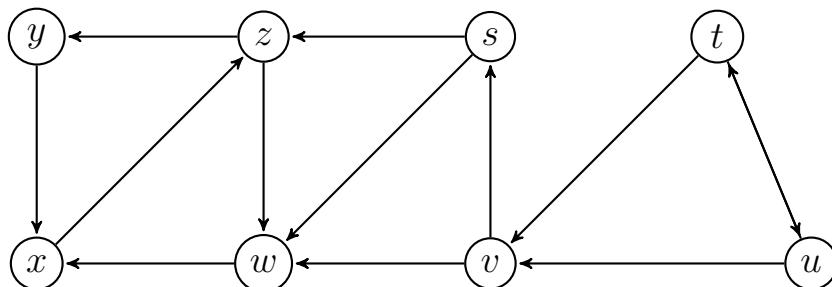


Figura 53: Esempio con visita in profondità

6.9 Classificazione

Naturalmente, dato un collegamento $u - v$, esistono $4! = 24$ permutazioni possibili della scoperta e dell'analisi di ciascun vertice, ma solamente 4 configurazioni sono compatibili con il teorema delle parentesi:

$$\begin{aligned} d[u] &< f[u] & d[v] &< f[v] \\ d[v] &< f[v] & d[u] &< f[u] \\ d[u] &< d[v] & f[v] &< f[u] \\ d[v] &< d[u] & f[u] &< f[v] \end{aligned}$$

In cui la prima configurazione è incompatibile con la presenza del collegamento $u - v$, in quanto prima di terminare l'analisi di u , ossia $f[u]$, è necessario esaminare tutta la sua adiacenza, giungendo inevitabilmente a scoprire v .

Allo stesso modo, se l'albero è orientato si può escludere anche la seconda configurazione; tuttavia, se l'albero non è orientato, tale configurazione rappresenta un ramo di taglio, dal momento che riguarda due alberi distinti, i quali non si parlano.

Nel terzo caso, si possono distinguere due casistiche possibili

- Nel primo caso, partendo da u , si scopre v tramite un cammino bianco;
- Nel secondo caso, semplicemente terminando v si procede a conoscere u partendo per vie traverse

Ovviamente, se l'albero è orientato la seconda opzione è impossibile, dal momento che se c'è il collegamento $u - v$, vi è anche $v - u$.

Nell'ultimo caso, invece, si ha che semplicemente v è un discendente di u per cui tornando indietro si completa l'analisi di u e anche quella di v , guardando semplicemente indietro. Quest'ultima configurazione segnala necessariamente la presenza di un ciclo, in quanto partendo da v , scoprendo u , si osserva nuovamente l'adiacenza di u , prima di terminare l'analisi di u e v ;

... continua ... 40:00 e oltre

non solo, ma se c'è un ciclo allora deve esserci necessariamente un ramo **backward**. Infatti, considerando un ciclo all'interno di un grafo:

immagine

Si supponga per ipotesi che all'istante $d[u]$ il vertice u sia il primo ad essere scoperto. Questo significa che esiste un cammino bianco che collega u a v ; ciò significa, per il teorema del cammin bianco, che anche se tale cammino viene impiegato o meno, implica che v sia un discendente di u ; pertanto, se all'istante $d[v]$ si scopre il nodo v , ad un certo istante si dovrà anche analizzare l'adiacenza di v e con questa anche il collegamento $v - u$, che è proprio il ramo *backwards* di cui si voleva dimostrare la presenza.

Osservazione: Pertanto, se il grafo è aciclico (DAG - Direct Acyced Graph), si ha sempre che $f[v] < f[u]$, in quanto prima si termina l'analisi dell'adiacenza di u e poi si analizza u .

6.10 Ordinamento topologico

All'interno di un grafo di propedeuticità, è fondamentale considerare le priorità di visita dei diversi nodi, seguendo la logica seguente:

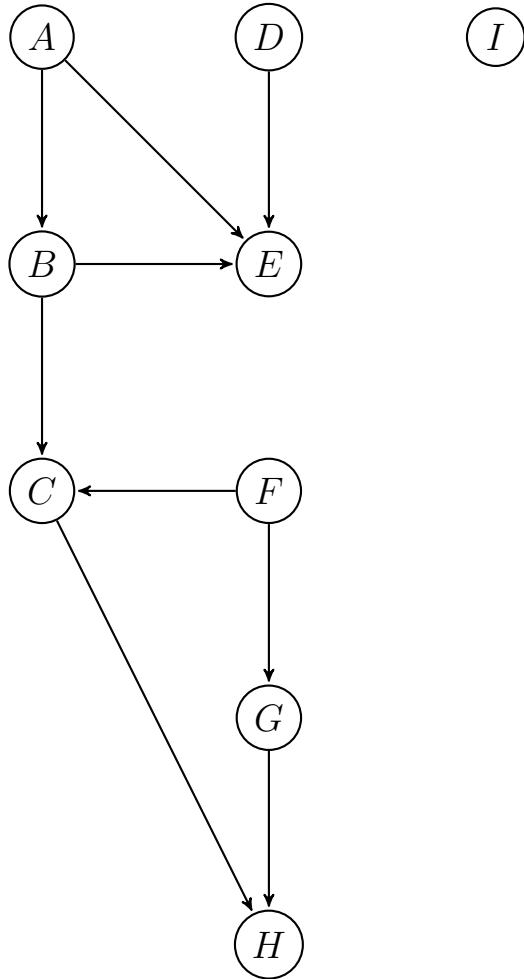


Figura 54: Esempio con visita in profondità

Allora, l'ordine di visita in profondità di tale grafo è dettato dall'ordine con cui viene descritto il grafo nella lista di adiacenza.

Pertanto, considerando a ritroso l'ordine di analisi dei vertici si ottiene esattamente una perfetta sequenza di ordinamento topologico che rispetta tutte le propedeuticità; pertanto, un effetto collaterale della visita in profondità è quella di ottenere l'ordinamento topologico del grafo: ciò funziona in quanto se si considera un DAG con un ramo $u - v$, allora è sempre vero che $f(v) < f(u)$.

Osservazione: Si osservi che come per la visita in ampiezza, anche la visita in profondità esegue le medesime operazioni, soltanto che il sequenziamento temporale è differente, per cui la complessità permane a $\Theta(|v| + |\xi|)$.

6.11 Algoritmo di Dijkstra

Per comprendere la potenza e la genialità dell'algoritmo di Dijkstra è fondamentale focalizzare l'attenzione sul concetto di **distanza**, la quale gode di 4 proprietà metriche fondamentali

1. $d(x, y) \geq 0$
2. $d(x, y) = 0$ se e solo se $x = y$
3. $d(x, y) = d(y, x)$
4. $d(x, z) < d(x, y) + d(y, z)$

Naturalmente è ovvio che è soddisfatta la seguente proprietà

$$0 \leq d(x, x) \leq \min [d(x, y), d(y, x)]$$

così come, dovendo passare per z e non essendoci simmetria nella definizione della distanza

$$d(x, z) + d(z, y) \geq d(x, y)$$

La distanza che viene impiegata all'interno dell'algoritmo di Dijkstra riguarda la configurazione particolare di un grafo in cui ogni **arco è pesato**, ossia presenta un costo $w \geq 0$: ciò significa che impiegare quel percorso comporta pagare un costo pari a w .

Alla luce di tale specifica, si possono descrivere le proprietà di tale *distanza*:

$$0 \leq d(x, x) \leq d(x, y)$$

$$0 \leq d(x, x) \leq d(y, x)$$

essendo le distanze (ossia i pesi) sempre positive. Inoltre, è anche vero che

$$d(x, y) = 0 \text{ anche se } x \neq y$$

che contrasta con le proprietà metriche. La conferma evidente che tale distanza non sia metrica, è il fatto che non vi è simmetria:

$$d(x, y) \neq d(y, x)$$

in quanto per raggiungere y a partire da x o viceversa vi possono essere percorsi diversi aventi pesi diversi: in ogni caso, la distanza tra x e y è la minima tra le distanze possibili.

Infine, la **disuguaglianza triangolare ordinata** vale necessariamente, in quanto si ha la somma di distanze necessariamente positive per raggiungere il vertice finale passando per vertici intermedi.

Lemma 6.1 *Se il cammino tra x e y è ottimo, allora anche i sottopercorsi intermedi sono necessariamente ottimi.*

13 Aprile 2022

I meccanismi di visita di un grafo noti sono due: visita in ampiezza e in profondità, ciascuno dei quali presentano un costo esecutivo paragonabile al mero inserimento dell'input, ovvero $\Theta(|\xi| + |v|)$. La visita in ampiezza presenta il pregio di calcolare le distanze dei nodi da un vertice **sorgente**: partendo da tale nodo, infatti, sarà sufficiente impiegare un solo albero per la descrizione del grafo, analizzando dapprima l'adiacenza della sorgente, e poi l'adiacenza di ogni nodo successivamente scoperto (impiegando la logica secondo la quale i vertici appena scoperti divengono grigi, mentre i vertici di cui è stata già analizzata l'adiacenza divengono neri).

In questo modo, naturalmente, è possibile capire la distanza di ogni nodo dalla radice, sempre nell'ipotesi che il peso di un arco sia sempre pari a 1 (a differenza dell'algoritmo di Dijkstra).

La visita in profondità, invece, prevede di considerare degli istanti temporali progressivi di visita e analisi: quando i vertici non sono corredati da un valore numerico sono bianchi, quando vi è un solo valore sono grigi, mentre se presentano due istanti di tempo allora avranno colore nero: ecco che in questo modo, quando da un vertice non si può più allontanarsi, tale vertice diviene nero e si deve tornare indietro per verificare una nuova possibilità di allontanamento.

È chiaro che si sarebbe potuto anche considerare un contatore pari al numero dei vertici del grafo e ogni volta che un nodo diviene nero si inserisce all'interno di un array nella posizione del contatore proprio tale vertice, in modo tale da ottenere un corretto ordine di visita del grafo (**topological sorting**). Naturalmente, se tale grafo è un DAG - Directed Acyclic Graph, ossia un grafo in assenza di cicli, tale grafo diviene un **grafo di propedeuticità**.

Osservazione: Nell'algoritmo di Dijkstra si parla sempre di distanze, ma in questo caso il costo, ossia il **weight**, di un arco non è unitario, ma è **sempre positivo**, sia in un verso che nell'altro (a differenza dell'algoritmo di Bellmann-Ford, computazionalmente più pesante, in cui i pesi possono essere anche negativi); viene da sé che se tutti i pesi sono identici, non ha senso impiegare l'algoritmo di Dijkstra, in quanto basta normalizzare tutti i pesi e ricondurli all'unità e impiegare la visita in ampiezza.

L'algoritmo di Dijkstra inizia con una procedura che prende il nome di **Initialize-Single-Source**, esposta di seguito:

Algorithm 33 Initialize-Single-Source

```
1: INITIALIZE-SINGLE-SOURCE( $\mathcal{G}, s$ )
2: for  $v \in V[\mathcal{G}]$ 
3:   do  $d[v] = +\infty$ 
4:    $\pi[v] = \text{NIL}$ 
5:  $d[s] = 0$ 
```

Dopodichè si impiega anche la procedura **Relax**:

Algorithm 34 Relax

```
1: RELAX( $u, v, w$ )
2: if  $d[v] > d[u] + w(u, v)$  then
3:    $d[v] = d[u] + w(u, v)$ 
4:    $\pi[v] = u$ 
```

in cui $w(u, v)$ permette di determinare il costo del collegamento tra i vertici u e v .

All'interno dell'algoritmo di Dijkstra si fa uso di un'operazione che prende il nome di EXTRACT-MIN(Q), che è un DEQUEUE prioritario che si potrebbe realizzare con un BUILD-HEAP che faccia rispettare un vincolo **min-heap** e non **max-heap**, semplicemente andando ad invertire le disuguaglianze presenti in BUILD-HEAP, che presenta il grande pregio di avere una complessità lineare.

Di seguito si epone il listato dell'algoritmo di Dijkstra, in cui w è la funzione che consente di determinare il costo del collegamento tra due vertici, mentre s è la sorgente di partenza:

Algorithm 35 Dijkstra

```

1: DIJKSTRA( $\mathcal{G}, w, s$ )
2: INITIALIZE-SINGLE-SOURCE( $\mathcal{G}, s$ )
3:  $S = \emptyset$ 
4:  $\mathcal{Q} = V[\mathcal{G}]$ 
5: while  $\mathcal{Q} \neq \emptyset$ 
6:   do EXTRACT-MIN( $\mathcal{Q}$ )
7:      $S = S \cup \{u\}$ 
8:     for  $v \in \text{Adj}[u]$ 
9:       do RELAX( $u, v, w$ )

```

Naturalmente, la coda \mathcal{Q} viene realizzata tramite un array, in quanto la dimensione massima è nota a priori, pari a al numero di vertici del grafo; analogamente, S è sempre un array di dimensione pari al numeri di vertici del grafo.

Esempio: Si consideri la descrizione iniziale del grafo seguente:

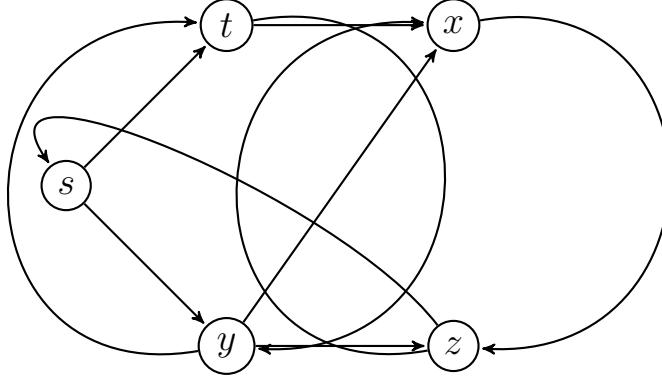


Figura 55: Esempio algoritmo Dijkstra

Tale algoritmo è corretto, tuttavia la dimostrazione della sua correttezza non è immediata ed evidente, così come la dimostrazione della sua complessità esecutiva.

Si espongono di seguito alcuni lemmi fondamentali per la dimostrazione della correttezza dell'algoritmo:

Lemma 6.2 Lemma della limitazione inferiore

Il cammino ottimo di collegamento tra la sorgente s e il vertice u è sempre minore o uguale della distanza $d[u]$, ovvero

$$d[u] \leq \delta(s, u)$$

Lemma 6.3 Lemma della convergenza

*Dato un cammino da s a v , in cui l'ultimo arco è $u - v$, che per ipotesi è un **cammino ottimo**, supponendo che $d[u] = \delta(u, v)$ a seguito del RELAX dell'arco $u - v$, allora anche v ha raggiunto il suo valore ottimale.*

Dimostrazione: A seguito del rilassamento di $u - v$, è sempre vero che

$$d[v] \leq d[u] + w(u, v)$$

ma è anche vero che

$$d[v] = \delta[u] + w(u, v)$$

ma per ipotesi il percorso $s - v$ è un percorso ottimo, per cui $d[v] = \delta[u] + w(u, v) = \delta[v]$. Il lemma della limitazione, però, afferma che $d[u] \geq \delta(v)$, per cui l'unica possibilità è che $d[v] = \delta[v]$.

Dimostrazione Dijkstra: In prima analisi, si consideri che l'algoritmo porta ad inserire nel contenitore S alcuni vertici che non dovrebbero starci; per esempio, se la sorgente $s \in S$ e un vertice $u \notin S$, potrebbe succedere che vi sia un collegamento $s - u$ che comprende vertici sia in S che non.

Allora, se $s - u$ è il percorso ottimo per collegare s a u , u è il primo vertice che viene inserito in S erroneamente.