

Ordinamenti e crescita della complessità

Informatica@SEFA 2018/2019 - Lezione 13

Massimo Lauria <massimo.lauria@uniroma1.it>*

Mercoledì, 31 Ottobre 2018

Nota bibliografica: Il contenuto di questa e di alcune delle prossime lezioni può essere trovato nei primi capitoli del libro *Introduzione agli Algoritmi e strutture dati* (Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein). Tuttavia il libro è grosso e costoso, quindi cerco di mettere tutto il materiale negli appunti del corso.

1 La misura della complessità computazionale

Quando consideriamo le prestazioni di un algoritmo non vale la pena considerare l'effettivo tempo di esecuzione o la quantità di RAM occupate. Queste cose dipendono da caratteristiche tecnologiche che niente hanno a che vedere con la qualità dell'algoritmo. Al contrario è utile considerare il numero di **operazioni elementari** che l'algoritmo esegue.

- accessi in memoria
- operazioni aritmetiche
- ecc. . .

Che cos'è un'operazione elementare? Per essere definiti, questi concetti vanno espressi su modelli computazionali **formali e astratti**, sui quali definire gli algoritmi che vogliamo misurare. Per noi tutto questo non è necessario. Ci basta capire a livello intuitivo il numero di operazioni elementari che i nostri algoritmi fanno per ogni istruzione.

*<http://massimolauria.net/courses/infosefa2018/>

1.1 Esempi

- una somma di due numeri float : 1 op.
- verificare se due numeri float sono uguali: 1 op.
- confronto lessicografico tra due stringhe di lunghezza n : n op.
- ricerca lineare in una lista Python di n elementi: n op.
- ricerca binaria in una lista ordinata di n elementi: $\log n$ op.
- inserimento nella 4a posizione in una lista di n elementi: n op.
- `seq[a:b]` : $b - a$ operazioni
- `seq[:]` : `len(seq)` operazioni
- `seq1 + seq2` : `len(seq1) + len(seq2)` operazioni

2 Un algoritmo di ordinamento: “Insertion sort”

Nella lezione precedente abbiamo visto due tipi di ricerca. Quella *sequenziale* o *lineare* e quella *binaria*. Per utilizzare la ricerca binaria è necessario avere la sequenza di dati ordinata. Oggi vedremo un **semplice** algoritmo di ordinamento.

- non è molto efficiente
- semplice da realizzare

L’idea dell’insertion sort è simile a quella di una persona che gioca a carte e che vuole ordinare la propria mano.

- Lascia la prima carta all’inizio.
- Poi prende la seconda carta e la mette in modo tale che la prima e seconda posizione siano ordinate.
- Poi prende la terza carta e la mette in modo tale che la prima e seconda e la terza posizione siano ordinate.
- ...

Al passo i -esimo si guarda l’ i -esima carta nella mano e la si **inserisce** tra le carte **già ordinate** nelle posizioni da 0 a $i - 1$.

In sostanza se una sequenza L ha n elementi, l'insertion sort fa iterazioni per i da 1 a $n - 1$. In ognuna:

- Garantisce che gli elementi nelle posizioni $0, \dots, i - 1$ siano ordinati;
- trova la posizione $j \leq i$ tale che $L[j-1] \leq L[i] < L[j]$;
- inserisce $L[i]$ al minimo nella posizione j , traslando gli elementi nella sequenza per fargli spazio.

Per prima cosa vediamo come trovare la posizione nella quale inserire un elemento x nella parte iniziale ordinata di una lista L , assumendo che questa parte sia ordinata. Se x è maggiore di tutti gli elementi della lista allora la funzione restituisce `None`.

```
def insertion_point(x, seq, end):
    for j in range(end+1):
        if seq[j] > x:
            return j
    return None
```

Poi scriviamo una funzione che sposta tutti di elementi in un intervallo della lista di una posizione verso destra.

```
def moveright(seq, start, end):
    for t in range(end, start, -1):
        seq[t] = seq[t-1]
```

```
def insertion_sort(seq):
    for i in range(1, len(seq)):
        value = seq[i]
        j = insertion_point(value, seq, i-1)
        if j is not None:
            moveright(seq, j, i)
            seq[j] = value
```

Vediamo un esempio con una sequenza arbitraria

```
dati = [21, -4, 3, 6, 1, -5, 19]
insertion_sort(dati)
```

```
Init      : [21, -4, 3, 6, 1, -5, 19]
Step i=1: [-4, 21, 3, 6, 1, -5, 19]
Step i=2: [-4, 3, 21, 6, 1, -5, 19]
Step i=3: [-4, 3, 6, 21, 1, -5, 19]
Step i=4: [-4, 1, 3, 6, 21, -5, 19]
Step i=5: [-5, -4, 1, 3, 6, 21, 19]
Step i=6: [-5, -4, 1, 3, 6, 19, 21]
```

Vediamo un esempio con una sequenza invertita

```
dati = [10,9,8,7,6,5,4,3,2,1]
insertion_sort(dati)
```

1
2

```
Init      : [10, 9, 8, 7, 6, 5, 4, 3, 2, 1]
Step i=1: [9, 10, 8, 7, 6, 5, 4, 3, 2, 1]
Step i=2: [8, 9, 10, 7, 6, 5, 4, 3, 2, 1]
Step i=3: [7, 8, 9, 10, 6, 5, 4, 3, 2, 1]
Step i=4: [6, 7, 8, 9, 10, 5, 4, 3, 2, 1]
Step i=5: [5, 6, 7, 8, 9, 10, 4, 3, 2, 1]
Step i=6: [4, 5, 6, 7, 8, 9, 10, 3, 2, 1]
Step i=7: [3, 4, 5, 6, 7, 8, 9, 10, 2, 1]
Step i=8: [2, 3, 4, 5, 6, 7, 8, 9, 10, 1]
Step i=9: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
```

2.1 Quante operazioni esegue l'insertion sort?

Per ogni ciclo i , con i da 1 a $n - 1$, l'algoritmo calcola la posizione j dove fare l'inserimento.

- Quando la funzione `insertion_point` restituisce `None` ha compiuto i operazioni di confronto, e non si deve effettuare nessuna traslazione.
- Quando la funzione `insertion_point` restituisce j ha compiuto $j + 1$ operazioni di confronto, e si devono effettuare $i - j - 1$ traslazioni.

Quindi al passo i -esimo si fanno circa i operazioni. Il totale è

$$\sum_{i=1}^{n-1} i = (n-1)n \approx n^2 \quad (1)$$

Esercizio: per quali input il numero di operazioni è maggiore? Per quali è minore?

2.2 Vale la pena ordinare prima di fare ricerche?

Abbiamo visto che t ricerche costano nt se si utilizza la ricerca sequenziale. E se vogliamo utilizzare **insertion sort** seguito da ricerche binarie?

Il costo è $n^2 + t \log n$, che è comparabile con nt solo se si effettuano almeno $t \approx n$ ricerche.

3 Notazione asintotica

Quando si contano il numero di operazioni in un algoritmo, non è necessario essere estremamente precisi. A che serve distinguere tra $10n$ operazioni e $2n$ operazioni se in un linguaggio di programmazione le $10n$ operazioni sono più veloci delle $2n$ operazioni implementate in un altro linguaggio? Al contrario è molto importante distinguere, ad esempio, $\frac{1}{100}n^2$ operazioni da $20\sqrt{n}$ operazioni. Anche una macchina più veloce paga un prezzo alto se esegue un algoritmo da $\frac{1}{100}n^2$ operazioni invece che uno da $20\sqrt{n}$.

Tuttavia un algoritmo asintoticamente più veloce spesso paga un prezzo più alto in fase di inizializzazione (per esempio perché deve sistemare i dati in una certa maniera). Dunque ha senso fare un confronto solo per lunghezze di input grandi abbastanza.

Per discutere in questi termini utilizziamo la notazione asintotica che è utile per indicare quanto cresce il numero di operazioni di un algoritmo, al crescere della lunghezza dell'input.

Definizione: date $f : \mathbb{N} \rightarrow \mathbb{R}$ e $g : \mathbb{N} \rightarrow \mathbb{R}$ diciamo che

$$f \in O(g)$$

se esistono $c > 0$ e n_0 tali che $0 \leq f(n) \leq cg(n)$ per ogni $n > n_0$.

- la costante c serve a ignorare il costo contingente delle singole operazioni elementari;
- n_0 serve a ignorare i valori piccoli di n .

Ad esempio $\frac{n^2}{100} + 3n - 10$ è in $O(n^2)$, non è in $O(n^\ell)$ per nessun $\ell < 2$ ed è in $O(n^\ell)$ per $\ell > 2$.

Complessità di un algoritmo: dato un algoritmo A consideriamo il numero di operazioni $t_A(x)$ eseguite da A sull'input $x \in \{0, 1\}^*$. Per ogni taglia di input n possiamo definire il **costo nel caso peggiore** come

$$c_A(n) := \max_{x \in \{0, 1\}^n} t_A(x) .$$

Diciamo che la complessità di un algoritmo A , ovvero il suo *tempo di esecuzione nel caso peggiore*, è $O(g)$ quando $c_A \in O(g)$. Per esempio abbiamo visto due algoritmi di ricerca (ricerca lineare e binaria) che hanno entrambi complessità $O(n)$, ed il secondo ha anche complessità $O(\log n)$.

La notazione $O(g)$ quindi serve a dare un limite approssimativo superiore al numero di operazioni che l'algoritmo esegue.

Esercizio: osservare che per ogni $1 < a < b$,

- $\log_a(n) \in O(\log_b(n))$; e
- $\log_b(n) \in O(\log_a(n))$.

Quindi specificare la base non serve.

Definizione (Ω e Θ): date $f : \mathbb{N} \rightarrow \mathbb{R}$ e $g : \mathbb{N} \rightarrow \mathbb{R}$ diciamo che

$$f \in \Omega(g)$$

se esistono $c > 0$ e n_0 tali che $f(n) \geq cg(n) \geq 0$ per ogni $n > n_0$. Diciamo anche che

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

se $f \in \Omega(g)$ e $f \in O(g)$ simultaneamente.

Complessità di un algoritmo (II): dato un algoritmo A consideriamo ancora il **costo nel caso peggiore** di A su input di taglia n come

$$c_A(n) := \max_{x \in \{0,1\}^n} t_A(x) .$$

Diciamo che la complessità di un algoritmo A , ovvero il suo *tempo di esecuzione nel caso peggiore*, è $\Omega(g)$ quando $c_A \in \Omega(g)$, e che ha complessità $\Theta(g)$ quando $c_A \in \Theta(g)$.

Per esempio abbiamo visto due algoritmi di ricerca (ricerca lineare e binaria) che hanno entrambi complessità $\Omega(\log n)$, ed il primo anche complessità $\Omega(n)$. Volendo essere più specifici possiamo dire che la ricerca lineare ha complessità $\Theta(n)$ e quella binaria ha complessità $\Theta(\log n)$.

Fate attenzione a questa differenza:

- se A ha complessità $O(g)$ allora A gestisce tutti gli input con al massimo g operazioni circa, asintoticamente;
- se A ha complessità $\Omega(g)$ allora per ogni n abbastanza grande, esiste un input di dimensione n per cui A richiede non meno di $g(n)$ operazioni circa.

3.1 Note su come misuriamo la lunghezza dell'input

Gli ordini di crescita sono il linguaggio che utilizziamo quando vogliamo parlare dell'uso di risorse computazionali, indicando come le risorse crescono in base alla dimensione dell'input.

La lunghezza dell'input n consiste nel numero di bit necessari a codificarlo. Tuttavia per problemi più specifici è utile, e a volte più comodo, indicare con n la dimensione "logica" dell'input. Ad esempio:

- trovare il massimo in una sequenza. n sarà la quantità di numeri nella sequenza;
- trovare una comunità in una rete sociale con x nodi, e y connessioni. La lunghezza dell'input può essere sia x che y a seconda dei casi;
- moltiplicare due matrici quadrate $n \times n$. La dimensione dell'input è il numero di righe (o colonne) delle due matrici.

Naturalmente misurare così la complessità ha senso solo se consideriamo, ad esempio, dati numerici di grandezza limitata. Ovvero numeri di grandezza tale che le operazioni su di essi possono essere considerate elementari senza che l'analisi dell'algoritmo ne sia compromessa.

4 Esercitarsi sulla notazione asintotica

Ci sono alcuni fatti ovvi riguardo la notazione asintotica, che seguono immediatamente le definizioni. Dimostrate che

1. per ogni $0 < a < b$, $n^a \in O(n^b)$, $n^b \notin O(n^a)$;
2. per ogni $0 < a < b$, $n^b \in \Omega(n^a)$, $n^a \notin \Omega(n^b)$;
3. se $f \in \Omega(h)$ e g una funzione positiva, allora $f \cdot g \in \Omega(h \cdot g)$;
4. se $f \in O(h)$ e g una funzione positiva, allora $f \cdot g \in O(h \cdot g)$;
5. che $(n + a)^b \in \Theta(n^b)$.
6. preso un qualunque polinomio $p = \sum_{i=0}^d \alpha_i x^i$ con $\alpha_d > 0$, si ha che $p(n) \in \Theta(n^d)$;
7. che $n! \in \Omega(2^n)$ ma che $n! \notin O(2^n)$;
8. che $\binom{n}{d} \in O(n^d)$;

- 9. che per ogni $a, b > 0$, $(\log(n))^a \in O(n^b)$;
- 10. che per ogni $a, b > 0$, $n^b \notin O((\log(n))^a)$;