

# Tempo di esecuzione e Analisi asintotica

26 febbraio 2019

# Algoritmi e pseudocodice

Un **algoritmo** è un procedimento per risolvere un **problema** mediante una sequenza finita di operazioni **elementari**.

Il procedimento deve essere descritto in maniera precisa e **univoca** allo scopo di essere eseguito automaticamente dall'**esecutore**.

Un **problema computazionale** è una relazione fra alcuni dati in ingresso (input) e alcuni dati in uscita (output)

L'algoritmo potrà essere implementato in un **programma** mediante un linguaggio di programmazione e una opportuna struttura dei dati.

# Obiettivi del corso

Abbiamo bisogno di algoritmi ogni qual volta vogliamo scrivere un programma.

**Obiettivo finale:** programmare in maniera **consapevole** e creativa

Dal problema reale al programma che lo risolve:

1. **Formalizzazione** del problema
2. Scelta della **tecnica** per progettare algoritmo
3. Scelta della **struttura dati** per organizzare i dati
4. Dimostrazione della **correttezza**
5. **Analisi** risorse usate / efficienza

# Corso di Progettazione di Algoritmi

Progettazione e analisi di algoritmi

- **Progettazione:** tecniche Divide-et-impera, Greedy, Programmazione dinamica, ricerca esaustiva intelligente
- **Analisi «asintotica»** delle risorse utilizzate: **spazio e tempo**.

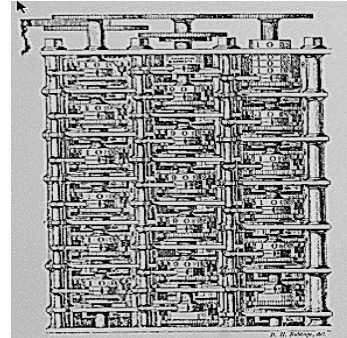
Nello studio dell'algoritmo (preventivo all'implementazione), esso è descritto tramite **pseudocodice** (o verbalmente)

# Computational Tractability

As soon as an Analytic Engine exists, it will necessarily guide the future course of the science. Whenever any result is sought by its aid, the question will arise - By what course of calculation can these results be arrived at by the machine in the shortest time? - *Charles Babbage*



Charles Babbage (1864)



Analytic Engine (schematic)

# Tempo di esecuzione

## Cosa significa?

*Esempio:* ricerca del massimo fra  $n$  numeri  $a_1, \dots, a_n$ .

Un algoritmo (in pseudo-codice):

```
max =  $a_1$ 
For i = 2 to n
    If ( $a_i > \text{max}$ ) then
        max =  $a_i$ 
    Endif
Endfor
```

Pseudocodice di [KT]

```
max  $\leftarrow a_1$ 
for i = 2 to n {
    if ( $a_i > \text{max}$ )
        max  $\leftarrow a_i$ 
}
```

Pseudocodice di [CLRS]

# Quale il tempo di esecuzione?

- Numero di secondi?
- Implementato con quale struttura dati, linguaggio, macchina, architettura, compilatore.....?
- Su quanti numeri? 100, 1.000, 1.000.000?

# Analisi

Studieremo il tempo in **funzione** della taglia **n** dell'input :

$$T(n)$$

Vogliamo analizzare l'efficienza dell'**algoritmo** (non di un programma) prima di scegliere implementazione, hardware etc.

**Cosa posso dire?**



# Tempo di esecuzione

**Tempo di esecuzione**  $T(n)$  sarà misurato in termini del **numero di operazioni elementari** per eseguire l'algoritmo su un input di taglia  $n$ .

Sono considerate **operazioni elementari** le operazioni che richiedono tempo **costante** (= non dipendente dalla taglia  $n$  dell'input).

**Per esempio:** assegnamento, incremento, confronto

Calcoleremo il tempo di esecuzione di un algoritmo, **ricorsivamente**, partendo dalla sua struttura.

# Struttura di un algoritmo (non ricorsivo)

Può essere costituito da:

- Una singola **istruzione elementare**
- Un **blocco** sequenziale di istruzioni (allo stesso livello di indentazione)
- Un'istruzione if: 

```
If(cond) then istr1  
                        Else istr2 Endif
```
- Un **ciclo** o loop for, while, repeat:
  - **For i=a to b {istr} Endfor**
  - **While (cond){istr}**
  - **Repeat {istr} until (cond)**

Pseudocodice di [KT]

# Struttura dell'esempio

```
1.  max = a1
2.  For i = 2 to n
3.      If (ai > max) then
4.          max = ai
5.      Endif
6.  Endfor
```

E' un **blocco** di 2 istruzioni:

1. Assegnamento

2-6. For

L'istruzione **For** delle linee 2-6 è:

**For i=a to b {istr0} Endfor**

dove **istr0** è l'istruzione if delle linee 3-5

L'istruzione **If** delle linee 3-5 è:

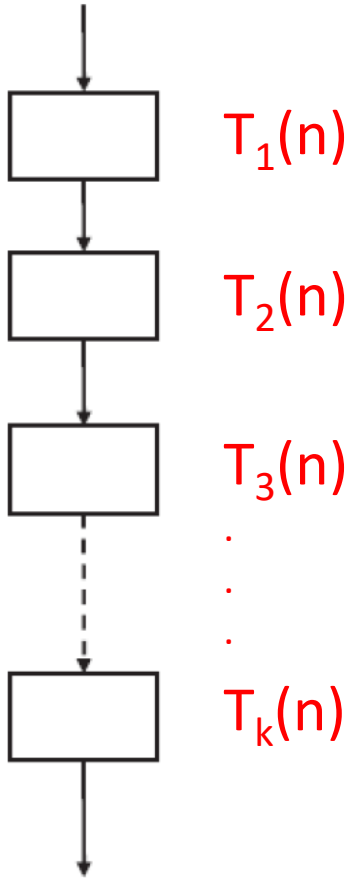
**If(cond) then istr1 Else istr2 Endif**

dove **cond** è il confronto: **a<sub>i</sub> > max**

**istr1** è l'assegnamento: **max = a<sub>i</sub>**

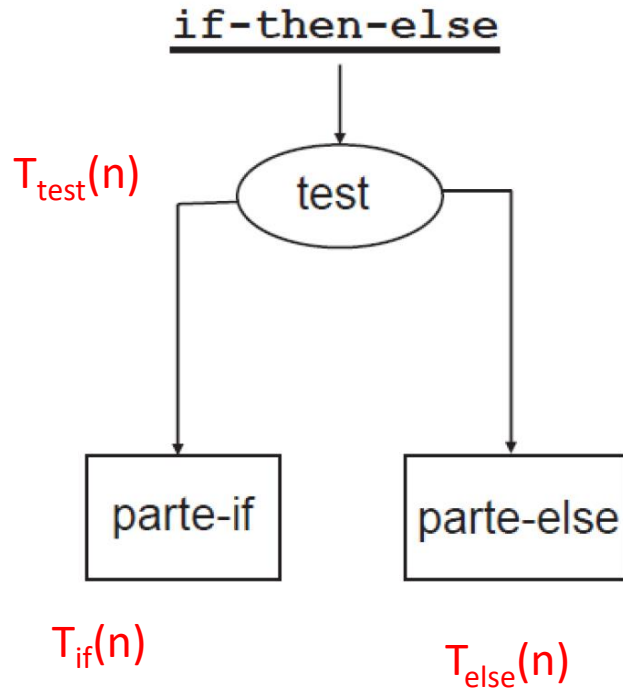
**istr2** è vuota

# Blocco sequenziale



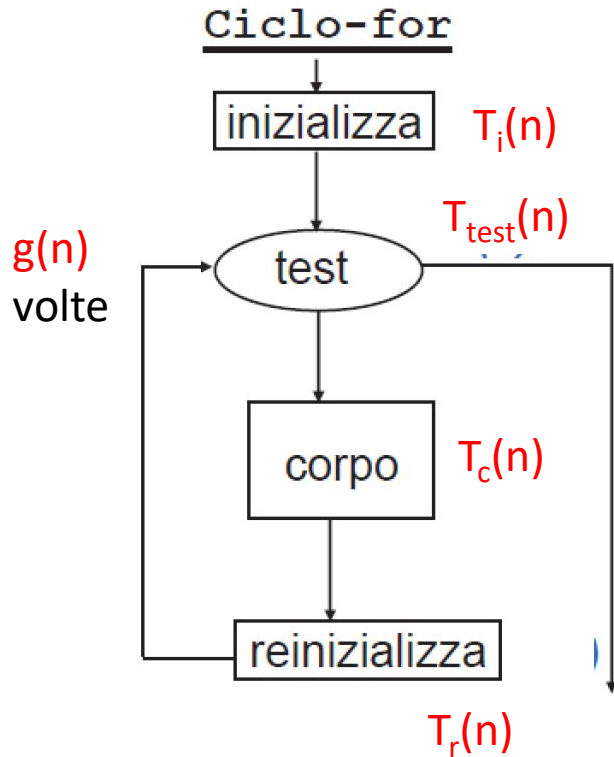
$$T(n) = T_1(n) + T_2(n) + T_3(n) + \dots + T_k(n)$$

# If then else



$$T(n) \leq T_{\text{test}}(n) + \max \{T_{\text{if}}(n), T_{\text{else}}(n)\}$$

# Ciclo for



$$T(n) = T_i(n) + T_{test}(n) + g(n) \times (T_{test}(n) + T_c(n) + T_r(n))$$

# Regole per il calcolo di $T(n)$ (caso non ricorsivo)

- il costo di istruzioni semplici, quali assegnazione, letture/scrittura di variabili é  $O(1)$
- il costo di un blocco sequenziale di istruzioni è pari alla somma dei costi delle singole istruzioni
- il costo di istruzioni tipo `if.... then ....else` é pari al tempo per effettuare il test (tipicamente  $O(1)$ ) piú  $O(\text{costo della alternativa piú costosa})$
- il costo di loop (`for`, `while`, `repeat`) é pari alla somma su tutte le iterazioni del costo di ogni iterazioni
- Usando queste regole si può ottenere la complessità di tempo di ogni algoritmo (ad eccezione degli algoritmi ricorsivi, che richiedono una tecnica diversa)

# Calcolo del tempo di esecuzione (1)

*Esempio:* algoritmo per la ricerca del massimo fra  $n$  numeri  $a_1, \dots, a_n$

```
1.  max = a1
2.  For i = 2 to n
3.      If (ai > max) then
4.          max = ai
5.      Endif
6.  Endfor
```

Pseudocodice di [KT]

Esecuzione del for (2-6):

n-1 {	i=2	2≤n?	if
	i=2+1=3	3≤n?	if
	i=3+1=4	4≤n?	if
	...	...	...
	i=n	n≤n?	if
	i=n+1	n+1≤n?	

Tempo di esecuzione = tempo esecuzione linea 1 + tempo esecuzione for(2-6)

Linea 1: 1 assegnamento

Linee 3-5 (una esecuzione): 1 confronto + 1 assegnamento eventualmente

Linee 2-6: 1 assegnamento (i=2) + (n-1) incrementi + n confronti + (n-1) if



# Calcolo del tempo di esecuzione (2)

```
1.  max = a1
2.  For i = 2 to n
3.      If (ai > max) then
4.          max = ai
5.      Endif
6.  Endfor
```

Taglia dell'input =  $n$

Tempo di un assegnamento =  $c_1$   
(costante = non dipende da  $n$ )

Tempo di un confronto =  $c_2$

Tempo di un incremento =  $c_3$

Linea 1: 1 assegnamento =  $c_1$

Linee 3 – 5 (una esecuzione): 1 confronto + 1 assegnamento eventualmente  $\leq c_2 + c_1$

Linea 2 - 6: 1 assegnamento +  $(n-1)$  incremento +  $n$  confronti +  $(n-1)$  if  $\leq$

$$\leq c_1 + (n-1) c_3 + n c_2 + (n-1) (c_2 + c_1)$$

$$T(n) \leq 2 c_1 + (n-1) c_3 + n c_2 + (n-1) (c_2 + c_1) = (c_3 + 2c_2 + c_1) n + c_1 - c_3 - c_2$$

# Calcolo del tempo di esecuzione (3)

*Esempio:* algoritmo per la ricerca del massimo fra  $n$  numeri  $a_1, \dots, a_n$

```
1.  max = a1
2.  For i = 2 to n
3.      If (ai > max) then
4.          max = ai
5.      Endif
6.  Endfor
```

Taglia dell'input =  $n$

Tempo di un assegnamento =  $c_1$   
(costante = non dipende da  $n$ )

Tempo di un confronto =  $c_2$

Tempo di un incremento =  $c_3$

Cosa posso dire adesso?

$$T(n) \leq 2 c_1 + (n-1) c_3 + n c_2 + (n-1) (c_2 + c_1) = (c_3 + 2c_2 + c_1) n + (c_1 - c_3 - c_2)$$

$$T(n) \leq A n + B$$

$T(n)$  è lineare!

dove  $A$  e  $B$  sono costanti non quantificabili a priori (dipendono dall'implementazione)

# Analisi asintotica di $T(n)$

Nell'analisi dell'**algoritmo** possiamo studiare  $T(n)$

- **Indipendentemente** dal valore delle **costanti** (costo di operazioni elementari)
- Al crescere della taglia dell'input

Questa si chiama **analisi asintotica** dove «asintotica» significa

- per  $n$  che tende a infinito
- per  $n$  arbitrariamente grande
- da un certo punto in poi
- per ogni  $n \geq n_0$

# Funzioni $T(n)$

Se  $T(n)$  rappresenta un tempo di esecuzione su un input di taglia  $n$ , allora:

$n$  è un intero positivo

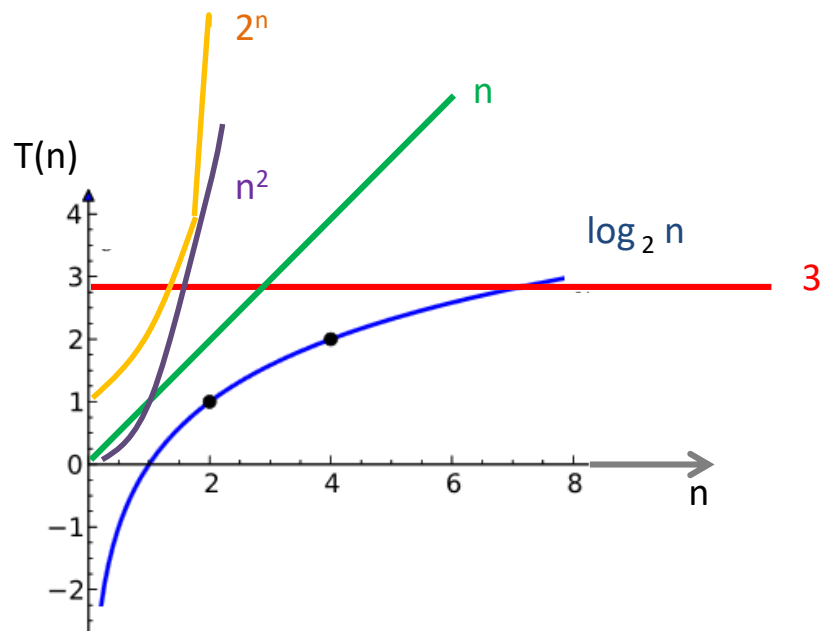
$T(n)$  è un reale positivo

$$T: \mathbb{N} \rightarrow \mathbb{R}_+$$

Inoltre  $T(n)$  è una funzione **non decrescente**

Ma vi sono vari modi di **crescita** della funzione  $T(n)$  al crescere di  $n$  (lineare, quadratica, polinomiale, esponenziale, ...): questa informazione può dirci già molto sull'**efficienza** di quella che sarà un'implementazione dell'algoritmo.

# Grafici delle funzioni



Approssimativamente....

# Funzioni standard

Confronteremo la crescita della funzione  $T(n)$  con alcune funzioni standard.

Fra le funzioni non decrescenti  $T: \mathbb{N} \rightarrow \mathbb{R}_+$  ci interesseranno principalmente le seguenti (e le loro combinazioni):

$$T_1(n) = c, \text{ con } c \text{ costante}$$

$$T_2(n) = \log n$$

$$T_3(n) = n$$

$$T_4(n) = n^2$$

$$T_5(n) = 2^n$$

# Notazioni asintotiche

Nell'analisi asintotica analizziamo  $T(n)$

1. A meno di costanti moltiplicative (perché non quantificabili)
2. Asintoticamente (per considerare input di taglia arbitrariamente grande)

Le notazioni asintotiche:  $O$ ,  $\Omega$ ,  $\Theta$

ci permetteranno il **confronto** tra funzioni, mantenendo queste caratteristiche

Diremo per esempio che l'algoritmo per la ricerca del massimo ha un tempo di esecuzione **lineare**,  $T(n) = O(n)$ , essendo  $T(n) \leq An + B$

# Vantaggi dell'analisi asintotica

- **Indipendente** da hardware
- Effettuabile con pseudocodice **prima** di implementare l'algoritmo
- Considera **infiniti** input

**Alternativa?** Analisi su dati campione.

*Svantaggi:* bisogna avere **già** implementato l'algoritmo;  
analizza numero **finito** di dati



# Casi di interesse

*Esempio:* problema dell'ordinamento

**INPUT:**  $n$  numeri  $a_1, \dots, a_n$

**OUTPUT:** permutazione dei numeri in cui ogni numero sia minore del successivo

Esistono svariati algoritmi che lo risolvono

**Qual è il tempo di esecuzione** per ordinare  $n$  elementi con un fissato algoritmo (per esempio InsertionSort)?

Anche per una stessa taglia  $n$  fissata, il tempo può dipendere dalla **distribuzione** dei numeri fra di loro

(es.: sono già ordinati, sono ordinati in senso inverso, sono tutti uguali, etc.)

# Caso peggiore, migliore, medio

Analisi del **caso peggiore**: qualunque sia la distribuzione dell'input,  $T(n)$  è limitata **superiormente** da  $f(n)$

Analisi del **caso migliore**: qualunque sia la distribuzione dell'input,  $T(n)$  è limitata **inferiormente** da  $g(n)$  (poco significativa)

Analisi del **caso medio**: nel caso di una distribuzione media o random (difficile da determinare)

# Worst-Case Analysis

- Worst case running time. Obtain bound on **largest possible** running time of algorithm on input of a given size  $N$ .
  - Generally captures efficiency in practice.
  - Draconian view, but hard to find effective alternative.
- Average case running time. Obtain bound on running time of algorithm on **random** input as a function of input size  $N$ .
  - Hard (or impossible) to accurately model real instances by random distributions.
  - Algorithm tuned for a certain distribution may perform poorly on other inputs.

Quando un algoritmo può essere considerato efficiente?

# Worst-Case Polynomial-Time

- Def. An algorithm is **efficient** if its running time is polynomial, i.e. **upper bounded** by a polynomial function (e.g.  $T(n) \leq A n^3 + B n + C$ )
- Justification: **It really works in practice!**
  - Although  $6.02 \times 10^{23} \times N^{20}$  is technically poly-time, it would be useless in practice.
  - In practice, the poly-time algorithms that people develop almost always have low constants and low exponents.
  - Breaking through the exponential barrier of brute force typically exposes some crucial structure of the problem.
- Exceptions.
  - Some poly-time algorithms do have high constants and/or exponents, and are useless in practice.
  - Some exponential-time (or worse) algorithms are widely used because the worst-case instances seem to be rare.

# Why It Matters

**Table 2.1** The running times (rounded up) of different algorithms on inputs of increasing size, for a processor performing a million high-level instructions per second. In cases where the running time exceeds  $10^{25}$  years, we simply record the algorithm as taking a very long time.

	$n$	$n \log_2 n$	$n^2$	$n^3$	$1.5^n$	$2^n$	$n!$
$n = 10$	< 1 sec	< 1 sec	< 1 sec	< 1 sec	< 1 sec	< 1 sec	4 sec
$n = 30$	< 1 sec	< 1 sec	< 1 sec	< 1 sec	< 1 sec	18 min	$10^{25}$ years
$n = 50$	< 1 sec	< 1 sec	< 1 sec	< 1 sec	11 min	36 years	very long
$n = 100$	< 1 sec	< 1 sec	< 1 sec	1 sec	12,892 years	$10^{17}$ years	very long
$n = 1,000$	< 1 sec	< 1 sec	1 sec	18 min	very long	very long	very long
$n = 10,000$	< 1 sec	< 1 sec	2 min	12 days	very long	very long	very long
$n = 100,000$	< 1 sec	2 sec	3 hours	32 years	very long	very long	very long
$n = 1,000,000$	1 sec	20 sec	12 days	31,710 years	very long	very long	very long

# Confronto efficienza

$n \log n$  vs  $2^n$

	$n$	$n \log_2 n$	$n^2$	$n^3$	$1.5^n$	$2^n$	$n!$
$n = 10$	< 1 sec	< 1 sec	< 1 sec	< 1 sec	< 1 sec	< 1 sec	4 sec
$n = 30$	< 1 sec	< 1 sec	< 1 sec	< 1 sec	< 1 sec	18 min	$10^{25}$ years
$n = 50$	< 1 sec	< 1 sec	< 1 sec	< 1 sec	11 min	36 years	very long
$n = 100$	< 1 sec	< 1 sec	< 1 sec	1 sec	12,892 years	$10^{17}$ years	very long
$n = 1,000$	< 1 sec	< 1 sec	1 sec	18 min	very long	very long	very long
$n = 10,000$	< 1 sec	< 1 sec	2 min	12 days	very long	very long	very long
$n = 100,000$	< 1 sec	2 sec	3 hours	32 years	very long	very long	very long
$n = 1,000,000$	1 sec	20 sec	12 days	31,710 years	very long	very long	very long

accettabile

non accettabile

# Efficiency = polynomial

---

Polynomial-time solvability emerged as a formal notion of efficiency by a gradual process, motivated by the work of a number of researchers including Cobham, Rabin, Edmonds, Hartmanis, and Stearns.

Similarly, the use of asymptotic order of growth notation to bound the running time of algorithms—as opposed to working out exact formulas with leading coefficients and lower-order terms—is a modeling decision that was quite non-obvious at the time it was introduced;



# Obiettivi formativi...

- Estrapolare il problema computazionale
- Riconoscere (analogie e variazioni con) problemi noti
- Scegliere tecnica/he più appropriata/e
- Applicarla/e correttamente
- Valutare l'efficienza

Nei prossimi esercizi si chiede di formalizzare il problema computazionale alla base del problema «reale» descritto.

# Appello 14 febbraio 2017

## Quesito 2 (23 punti) (*San Valentino*)

Oggi è San Valentino e volete comprare un regalo per il vostro partner. Andate nel suo negozio preferito e selezionate un insieme  $S$  di  $n$  oggetti che sicuramente sarebbero graditi. Ma, guardando nel vostro portafoglio, vi accorgete che con la somma che avete, di certo non potete comprarli tutti! Decidete allora di assegnare ad ogni oggetto il presunto valore di gradimento e di selezionare quindi un insieme di oggetti di  $S$  che abbiano un valore di gradimento massimo, ma che non superi la somma che avete a disposizione. Il problema non è però di facile soluzione, anche perché non avete molto tempo a disposizione, in quanto il negozio sta per chiudere.

Formalizzate il problema reale in un problema computazionale e risolverlo nel modo più efficiente possibile con la tecnica che ritenete più opportuna. E' necessario descrivere l'algoritmo soluzione e valutarne l'efficienza.

# Appello 15 gennaio 2015

## **Quesito 2** (24 punti)

Dopo le festività natalizie avete messo su qualche chilo, e decidete di cominciare una dieta. Ne avete trovato una che fa al caso vostro.

La Dieta consiste nel comporre ogni pasto scegliendo fra alcuni cibi che avete a disposizione, ognuno con un assegnato numero di calorie, senza superare una data quantità di calorie a pasto. Potete però scegliere secondo il vostro gusto. Ad ogni porzione di cibo a disposizione, assegnate quindi un vostro personale "grado di appetibilità" (un intero da 1 a 10).

Il problema è allora di calcolare il pasto che non superi il massimo di calorie ammesse in quel pasto, ma che sia il più appetitoso possibile. Ricordate che ogni porzione di cibo va scelta per intera al più una volta.

Formalizzate il problema reale in un problema computazionale. e risolverlo con la tecnica che ritenete più opportuna affinché la soluzione sia il più efficiente possibile. Giustificate le risposte.

# Appello 29 gennaio 2015

## **Quesito 2** (24 punti)

Dopo la Laurea in Informatica avete aperto un campo di calcetto che ha tantissime richieste e siete diventati ricchissimi. Ciò nonostante volete guadagnare sempre di più, per cui avete organizzato una sorta di asta: chiunque volesse affittare il vostro campo (purtroppo è uno solo), oltre ad indicare da che ora a che ora lo vorrebbe utilizzare, deve dire anche quanto sia disposto a pagare. Il vostro problema è quindi scegliere le richieste compatibili per orario, che vi diano il guadagno totale maggiore. Avete trovato una soluzione, ma è molto lenta. Vi ricordate allora che al corso di Algoritmi vi avevano tempestato con le varie tecniche di progettazione di algoritmi: potranno esservi utili (almeno una volta nella vita)?

Formalizzate il problema reale in un problema computazionale. e risolverlo in maniera che sia il più efficiente possibile. Giustificate le risposte

# Appello 19 febbraio 2015

## **Quesito 2** (24 punti)

Dopo la Laurea in Informatica avete aperto un campo di tennis, ma gli affari non vanno benissimo. Decidete di proporre un'offerta: chi vuole affittare il campo, indipendentemente dalla durata del noleggio, paga sempre lo stesso prezzo fisso di  $x$  euro. Sarete però voi a decidere quali richieste accontentare, dopo aver raccolto tutte le richieste, con i rispettivi orari di inizio e di fine, per una stessa giornata.

Il vostro problema è di scegliere le richieste compatibili per orario, che vi diano il massimo guadagno totale.

Formalizzate il problema reale in un problema computazionale. e risolverlo in maniera che sia il più efficiente possibile, per quanto riguarda sia il tempo di esecuzione che lo spazio di memoria necessario.

Giustificate la correttezza della vostra soluzione.

# Appello 17 aprile 2015

## **Quesito 2** (24 punti)

E' domenica pomeriggio, piove e non avete voglia di studiare (né di vedere le partite!).

Coi vostri amici decidete di fare una “maratona” di film. Potete scegliere fra un certo numero di film, di cui conoscete la durata esatta, e volete vedere il maggior numero di film possibili nelle ore che avete a disposizione. Il vostro problema è ora di selezionare i film. Ognuno di voi ha un'idea differente, dalla più semplice alla più complicata, ma nessuno riesce a convincere gli altri che la propria soluzione è la migliore. E il tempo passa. Vi ricordate allora che al corso di Algoritmi vi avevano tempestato con le varie tecniche di progettazione di algoritmi, sui modi per dimostrarne la correttezza e valutarne la bontà: potranno esservi utili (almeno una volta nella vita)?

Formalizzate il problema reale in un problema computazionale. e risolverlo nella maniera più efficiente possibile e in modo che, soprattutto, riusciate a convincere gli altri che la vostra soluzione funziona perfettamente. Giustificate le risposte.

## DEFINIRE PROBLEMA COMPUTAZIONALE

### **Quesito** (22 punti) (*Campi di calcetto*)

Dopo il successo del vostro primo campo di calcetto, avete aperto molti altri campi di calcetto, all'interno di un unico complesso. Ogni giorno raccogliete le richieste per utilizzare i vostri campi, ognuna specificata da un orario di inizio e un orario di fine. Oramai avete un numero di campi sufficiente ad accontentare sempre tutte le richieste. Volete però organizzare le partite nei campi in modo da accontentare tutti, senza che vi siano sovrapposizioni di orari, ma con il minimo numero possibile di campi (la manutenzione costa!).

Formalizzate il problema reale in un problema computazionale.