Qual è il minimo numero di confronti che garantisca il merge di 2 liste ordinate?

In seno all'algoritmo per l'ordinamento chiamato *MergeSort* abbiamo incontrato una sottoprocedura, chiamata *Merge*, il cui ruolo era quello di combinare in un'unica lista ordinata gli $n+m$ elementi di due liste (di n ed m elementi rispettivamente) ricevute in input. Il numero di confronti impiegato dalla semplice procedura che avevamo preso in considerazione era $n+m-1$, nel caso peggiore.

Domanda 1 *Poteva la nostra Merge impiegare meno di $n+m-1$ confronti, a seconda delle sequenze ricevute in input?*

Domanda 2 *Sapresti dire il numero di confronti impiegato dalla nostra Merge in media?*

Domanda 3 *Quale è il minimo numero di confronti che un qualsiasi algoritmo dovrà impiegare nel suo caso peggiore per fare il merging di due liste ordinate?*

Nel seguito, assumeremo di ricevere in input due liste ordinate

$$A_1 < A_2 < \ldots < A_m$$

e

$$B_1 < B_2 < \ldots < B_n$$

ed assumeremo per semplicità che tutti gli elementi siano distinti. Indicheremo con $M(n,m)$ il minimo numero di confronti che risultino sempre sufficienti alla fusione delle due liste in inputs in un'unica lista ordinata.

Domanda 4 *Sapresti dire perchè sappiamo già che $M(n,m) \leq m+n-1$? Riesci a giustificarti il significato della frase seguente: "del fatto che $M(n,m) \leq m+n-1$ abbiamo già dato una dimostrazione algoritmica".*

Chiaramente, gli elementi A potranno apparire inframmezzati agli elementi B in un totale di $\binom{m+n}{m}$ modi diversi. Pertanto, poichè l'esito di un confronto mi consegna al più un bit di informazione, ne consegue che

$$M(n,m) \geq \left\lceil \log_2 \binom{m+n}{m} \right\rceil \tag{1}$$

Esercizio 1 *Si noti l'eleganza con cui il lower bound 1 è stato ottenuto sulla base di semplici considerazioni attinenti alla quantità di informazione, poggiando sull'ovvio principio che l'informazione non possa venire dal nulla. Sapresti formalizzare una nozione di "quantità di informazione" che ti consenta di acquisire maggiore fiducia nella dimostrazione ora suggerita? Prova a derivare, in base alla stessa tipologia di argomenti, il lower bound di $n \log n - O(n)$ confronti dato per l'ordinamento.*

Esercizio 2 *Condurre un'analisi asintotica sul lower bound 1 assumendo che m sia costante e facendo crescere n. Il tuo obbiettivo nel condurre tale analisi sarà quello di evidenziare quanto esso si discosti dall'upper bound $M(n,m) \leq m+n-1$.*

Esercizio 3 *Se sarai riuscito a portare sufficientemente avanti l'analisi asintotica di cui all'esercizio 2, avrai realizzato che il lower bound 1 lascia ampio spazio al miglioramento della procedura Merge. Si mostri come sia in effetti possibile ridurre il numero di confronti nel caso peggiore quando n cresce ed m è invece costante.*

Ma anche nel caso per noi più significativo, quando cioè $m = n$, il lower bound proposto non consente di farci un sandwich con $M(n,n) \leq 2n-1$ e mostarda, al fine di argomentare l'ottimalità della nostra procedura *Merge*.

$$M(n,n) \geq \left\lceil \log_2 \binom{2n}{n} \right\rceil = \left\lceil \log_2 \frac{(2n)!}{(n!)(n!)} \right\rceil = \left\lceil \log_2 \frac{2n}{n} \frac{2n-1}{n-1} \cdots \frac{n+1}{1} \right\rceil$$

Esercizio 4 *Partire dalla diseguaglianza*

$$M(n,n) \geq \left\lceil \sum_{i=1}^{n} \log_2 \frac{n+i}{i} \right\rceil$$

praticamente ricavata qui sopra per giungere alla formula asintotica

$$M(n,n) \geq 2n - \frac{1}{2}\log n + O(1)$$

Pertanto, il lower bound basato su considerazioni attinenti alla quantità di informazione, ha il corretto ordine di crescita (ed anche la corretta costante) ma non consente di dimostrare l'ottimalità della nostra procedura di *Merge*. E tuttavia, è possibile dimostrare $M(n,n) = 2n - 1$.

Teorema 1 $M(n,n) = 2n - 1$, *for* $n \geq 1$.

Proof: Si consideri il caso in cui la giusta soluzione al problema del merging sia costituita da

$$B_1 < A_1 < B_2 < A_2 < \cdots < B_n < A_n$$

Ci si chieda:

- sarebbe stato sicuro uscire dalla procedura senza aver eseguito un confronto del tipo $A_i \leftrightarrow B_i$?
- sarebbe stato sicuro uscire dalla procedura senza aver eseguito un confronto del tipo $A_i \leftrightarrow B_{i+1}$?

Se ne traggano le debite conclusioni. ∎

Per esercizio, dimostrare ora il seguente teorema, anch'esso rilevante per le applicazioni di *Merge* in seno all'algoritmo *MergeSort*.

Teorema 2 $M(n, n+1) = 2n$, *for* $n \geq 0$.