Analisi e progetto di algoritmi
Esercizi di programmazione dinamica

Notazione

Le sequenze di numeri interi vengono indicate con le lettere maiuscole, ad esempio $S = \langle 1, 7, 4, 9 \rangle$. L'$i$-esimo prefisso di una sequenza $X = \langle x_1, x_2, \ldots, x_n \rangle$ viene indicato con $X_i = \langle x_1, x_2, \ldots, x_i \rangle$, ad esempio considerata la sequenza S definita precedentemente, il prefisso di indice 2 è $S_2 = \langle 1, 7 \rangle$. Secondo questa definizione risulta evidente che $X_0 = \epsilon$, ossia la sequenza vuota. È definita sottosequenza una qualsiasi successione di elementi (non necessariamente contigui) appartenenti alla medesima sequenza. Ad esempio $\langle 7, 4 \rangle$ e $\langle 7, 9 \rangle$ sono sottosequenze di S, mentre $\langle 7, 9, 4 \rangle$ non è una sottosequenza di S.

Ogni esercizio presenta la seguente struttura:

1. definizione del problema principale `PB` e di una sua versione ridotta `PBR`
2. definizione dei sottoproblemi comuni e delle variabili descrittive
3. *possibile introduzione di un problema ausiliario e relativi sottoproblemi*
4. calcolo delle soluzioni dei sottoproblemi tramite equazioni di ricorrenza (passo ricorsivo e caso base)
5. definizione della soluzione del problema (eventualmente in funzione delle soluzioni dei sottoproblemi del problema ausiliario)
6. presentazione dello pseudocodice e della complessità dell'algoritmo

1 Più lunga sottosequenza comune (LCS)

1.1 Definizione del problema

Il problema `PB` è definito come segue:

date due sequenze X e Y, rispettivamente di m ed n numeri interi, si determini ***UNA*** *tra le più lunghe sottosequenze comuni a X e Y*

Ad esempio, date $X = \langle 2, 4, 7, 11, 21, 14, 1 \rangle$ e $Y = \langle 2, 7, 4, 23, 21, 14, 1, 8 \rangle$, una tra le più lunghe sottosequenze comuni è la sottosequenza $\langle 2, 4, 21, 14, 1 \rangle$.

Quello che si andrà a risolvere, comunque, è una versione ridotta del problema `PB`. La versione ridotta `PBR` è definita come segue:

date due sequenze X e Y, rispettivamente di m ed n numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze comuni a X e Y

Ad esempio, date $X = \langle 2, 4, 7, 11, 21, 14, 1 \rangle$ e $Y = \langle 2, 7, 4, 23, 21, 14, 1, 8 \rangle$, la lunghezza di una tra le più lunghe sottosequenze comuni è 5.

1.2 Sottoproblemi e variabili associate

Il problema `PBR` contiene diversi sottoproblemi ognuno dei quali non ha come input la coppia (X, Y) ma una coppia di prefissi di tali sequenze.

Il **sottoproblema di dimensione** (i, j) è definito come segue:

date due sequenze X e Y, rispettivamente di m ed n numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze comuni al prefisso X_i e al prefisso Y_j

Dato che $0 \leq i \leq m$ e $0 \leq j \leq n$, si ottengono $(m+1) \cdot (n+1)$ sottoproblemi (i e j possono valere 0 in quanto si deve considerare anche il caso in cui un prefisso sia la sequenza vuota). Ad ogni sottoproblema di `PBR` è associata una **variabile**.

Considerato il sottoproblema di dimensione (i, j), la variabile ad esso associata è $c_{i,j}$ ed è così definita:

$$c_{i,j} = \textit{lunghezza di una tra le più lunghe sottosequenze comuni a } X_i \textit{ e } Y_j$$

Per determinare la soluzione di un qualsiasi sottoproblema di dimensione (i, j), oltre all'input del problema, si utilizzeranno le soluzioni dei sottoproblemi di dimensione minore. Per esempio, dato il sottoproblema di dimensione $(2, 2)$, il valore della variabile $c_{2,2}$ si otterrà utilizzando le variabili $c_{0,0}$, $c_{0,1}$, $c_{1,0}$, $c_{1,1}$, $c_{1,2}$ e $c_{2,1}$. Si noti, però, che ognuna di queste variabili è da considerare come una ***black-box***: si può utilizzare ma non è possibile conoscerne il contenuto.

Considerato un qualunque sottoproblema di dimensione (i, j), assumendo di aver già risolto tutti i sottoproblemi di dimensione minore (ossia sapendo di poter utilizzare le variabili $\{c_{0,0}, \ldots, c_{i-1,j}, c_{i,j-1}\}$ senza però sapere effettivamente quanto valgono), come è possibile ottenere la soluzione $c_{i,j}$ andando a combinare le soluzioni dei sottoproblemi di dimensione minore a (i, j) ed, eventualmente, accodando gli elementi x_i e y_j estratti dall'input (X_i, Y_i)?

1.3 Equazioni di ricorrenza

Per rispondere a questa domanda, si devono andare a definire le **equazioni di ricorrenza**.

Le equazioni di ricorrenza si scrivono senza sapere il valore delle soluzioni dei sottoproblemi ma sapendo solamente che tali soluzioni esistono e si possono utilizzare

Un'equazione di ricorrenza è composta da:

- un **caso base** che definisce i casi più semplici che possono essere subito risolti senza ricorrere alle soluzioni dei sottoproblemi più piccoli

- un **passo ricorsivo** che definisce come risolvere i casi più complessi a partire dalle soluzioni dei sottoproblemi più piccoli (che si assume essere stati già risolti)

1.3.1 Caso base: (i, j) con $i = 0 \vee j = 0$

Il caso base si ha per un qualunque sottoproblema di dimensione (i, j) con $i = 0 \vee j = 0$, ossia quando uno dei due prefissi considerati è la sequenza vuota. In questo caso, è facile ottenere il valore della variabile $c_{i,j}$ in quanto la lunghezza di una tra le più lunghe sottosequenze comuni fra una qualsiasi sequenza e la sequenza vuota è 0 (ossia la lunghezza della sequenza vuota). Per questa ragione, il caso base è scrivibile come:

$$c_{i,j} = 0 \quad \textit{se } i = 0 \vee j = 0$$

1.3.2 Passo ricorsivo: (i, j) con $i > 0$ e $j > 0$

Il passo ricorsivo si ha per un qualunque sottoproblema di dimensione (i, j) con $i > 0 \ \wedge \ j > 0$, ossia quando si vanno a considerare due prefissi $X_i = \langle x_0, x_1, ..., x_i \rangle$ e $Y_j = \langle y_0, y_1, ..., y_j \rangle$ entrambi diversi dalla sequenza vuota. I dati disponibili per calcolare $c_{i,j}$ sono: l'input X ed in particolare l'elemento x_i, l'input Y ed in particolare l'elemento y_j e tutte le variabili $\{c_{0,0}, \ldots, c_{i-1,j}, c_{i,j-1}\}$. Per calcolare $c_{i,j}$ e quindi risolvere il problema di dimensione (i, j) è necessario considerare i due seguenti casi:

- $x_i = y_j$: se i due elementi considerati sono identici allora la lunghezza della più lunga sottosequenza comune fra X_i e Y_j è uguale alla lunghezza della più lunga sottosequenza comune fra X_{i-1} e Y_{j-1} (ossia il valore di $c_{i-1,j-1}$) aumentata di uno (l'elemento comune x_i è accodato ad una più lunga sottosequenza comune a X_{i-1} e Y_{j-1} correlata al sottoproblema di dimensione $(i-1, j-1)$ e che ha quindi ha lunghezza $c_{i-1,j-1}$). Il tutto è esemplificato dalla Figura 1 e può essere scritto come:

$$c_{i,j} = 1 + c_{i-1,j-1} \qquad \textit{se } x_i = y_j$$

- $x_i \neq y_j$: se i due elementi considerati sono differenti, allora la lunghezza della più lunga sottosequenza comune è data dalla soluzione di uno dei sottoproblemi di dimensione minore. Siccome gli elementi finali dei due prefissi considerati sono differenti, non possono appartenere entrambi alla soluzione: si vanno a considerare le soluzioni $c_{i-1,j}$ (ossia la soluzione del sottoproblema di dimensione $(i-1, j)$ con input il prefisso X_{i-1} e il prefisso Y_j: si è tolto da X_i l'elemento x_i) e $c_{i,j-1}$ (ossia la soluzione del sottoproblema di dimensione $(i, j-1)$ con input il prefisso X_i e il prefisso Y_{j-1}: si è tolto da Y_j l'elemento y_j) per poi andare a selezionare il massimo fra questi due valori. Il tutto può essere scritto come:

$$c_{i,j} = \max\{c_{i-1,j}, c_{i,j-1}\} \qquad \textit{se } x_i \neq y_j$$

Il passo ricorsivo è quindi scrivibile come:

$$c_{i,j} = \begin{cases} 1 + c_{i-1,j-1} & \text{se } x_i = y_j \\ \max\{c_{i,j-1}, c_{i-1,j}\} & \text{se } x_i \neq y_j \end{cases}$$

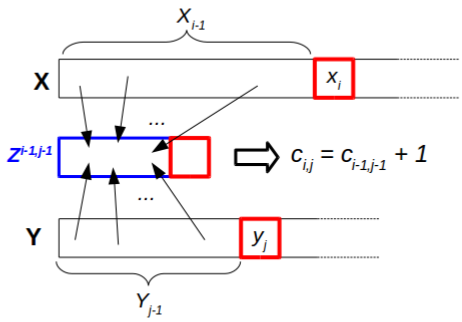

Figura 1: Sottosequenza comune più lunga: passo ricorsivo nel caso in cui $x_i = y_j$. L'elemento x_i viene accodato a $Z^{i-1,j-1}$ che è una più lunga sottosequenza comune a X_{i-1} e Y_{j-1}

1.4 Soluzione del problema

Una volta calcolati i valori $c_{0,0}, c_{0,1}, \ldots, c_{m,n}$ si hanno a disposizione tutte le lunghezze delle sottosequenze comuni massimali fra qualsiasi prefisso di X e qualsiasi prefisso di Y. La soluzione al problema `PBR` è:

$$c_{m,n}$$

Esempio Date $X = \langle 2, 4, 7, 11, 21, 14, 1 \rangle$ e $Y = \langle 2, 7, 4, 23, 21, 14, 1, 8 \rangle$, tutti i valori $c_{i,j}$ relativi sono contenuti nella matrice presentata in Figura 2.

1.5 Implementazione

L'algoritmo 1 presenta lo pseudocodice che calcola tutti i valori $c_{0,0}, c_{0,1}, \ldots, c_{m,n}$ basandosi sulle equazioni di ricorrenza definite nella Sezione 1.3. Anziché utilizzare la ricorsione pura, tutti i valori sono calcolati con tecnica **bottom-up** in modo tale da risolvere ogni sottoproblema una volta sola. Questo permette di risolvere il problema in $\mathcal{O}(m \cdot n)$ occupando $\Theta(m \cdot n)$ spazio in memoria (ossia una matrice che contiene i vari $c_{i,j}$).

2 Più lunga sottosequenza crescente (LGS)

2.1 Definizione del problema

Il problema `PB` è definito come segue:

*data una sequenza X di m numeri interi, si determini **UNA** tra le più lunghe sottosequenze crescenti di X*

Ad esempio, data $X = \langle 2, 4, 7, 6, 11, 13, 21, 14, 1 \rangle$, una tra le più lunghe sottosequenze crescenti di X è la sottosequenza $\langle 2, 4, 6, 11, 13, 21 \rangle$.

Quello che si andrà a risolvere, comunque, è una versione ridotta del problema `PB`. La versione ridotta `PBR` è definita come segue:

		0	1	2	3	4	5	6	7	8	***j***
		ε	2	7	4	23	21	14	1	8	y_j
0	ε	0	0	0	0	0	0	0	0	0	
1	2	0	1	1	1	1	1	1	1	1	
2	4	0	1	1	2	2	2	2	2	2	
3	7	0	1	2	2	2	2	2	2	2	
4	11	0	1	2	2	2	2	2	2	2	
5	21	0	1	2	2	2	3	3	3	3	
6	14	0	1	2	2	2	3	4	4	4	
7	1	0	1	2	2	2	3	4	5	5	
i	x_i										

Figura 2: Sottosequenza comune più lunga: matrice di tutti i $c_{i,j}$ ottenuti considerando le sequenze $X = \langle 2, 4, 7, 11, 21, 14, 1 \rangle$ e $Y = \langle 2, 7, 4, 23, 21, 14, 1, 8 \rangle$. In rosso è evidenziata la soluzione del problema ridotto

---

Algorithm 1 Lunghezza di una più lunga sottosequenza comune fra due sequenze X e Y di lunghezza m e n rispettivamente.

---

procedure LCS(X, Y)
  for $i \leftarrow 0$ **to** m **do**
    $c[i, 0] \leftarrow 0$
  for $j \leftarrow 0$ **to** n **do**
    $c[0, j] \leftarrow 0$
  for $i \leftarrow 1$ **to** m **do**
    for $j \leftarrow 1$ **to** n **do**
      if $x_i = y_j$ **then**
        $c[i, j] \leftarrow 1 + c[i-1, j-1]$
      else
        $c[i, j] \leftarrow max\{c[i, j-1], c[i-1, j]\}$
  return $c[m, n]$

---

data una sequenza X di m numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze crescenti di X

Ad esempio, data $X = \langle 2, 4, 7, 6, 11, 13, 21, 14, 1 \rangle$, la lunghezza di una tra le più lunghe sottosequenze crescenti di X è 6.

2.2 Sottoproblemi e variabili associate

Il problema `PBR` contiene diversi sottoproblemi ognuno dei quali non ha come input la sequenza X ma un suo prefisso.

Il **sottoproblema di dimensione** (i) è definito come segue:

data una sequenza X di m numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze crescenti di X_i

Dato che $0 \leq i \leq m$, si ottengono $(m+1)$ sottoproblemi (i può valere 0 in quanto si deve considerare anche il caso in cui il prefisso sia la sequenza vuota). Ad ogni sottoproblema di `PBR` è associata una **variabile**.

Considerato il sottoproblema di dimensione (i), la variabile ad esso associata è c_i ed è così definita:

$$c_i = \textit{lunghezza di una tra le più lunghe sottosequenze crescenti di } X_i$$

Per determinare la soluzione di un qualsiasi sottoproblema di dimensione (i), oltre all'input del problema, si utilizzeranno le soluzioni dei sottoproblemi di dimensione minore. Per esempio, dato il sottoproblema di dimensione (3), il valore della variabile c_3 si otterrà utilizzando le variabili c_0, c_1 e c_2. Si noti, però, che ognuna di queste variabili è da considerare come una ***black-box***: si può utilizzare ma non è possibile conoscerne il contenuto.

Considerato un qualunque sottoproblema di dimensione (i), assumendo di aver già risolto tutti i sottoproblemi di dimensione minore (ossia sapendo di poter utilizzare le variabili $\{c_0, \ldots, c_{i-1}\}$ senza però sapere effettivamente quanto valgono), come è possibile ottenere la soluzione c_i andando a combinare le soluzioni dei sottoproblemi di dimensione minore a (i) ed, eventualmente, accodando l'elemento x_i estratto dall'input X_i?

Purtroppo, però, il problema così definito non è risolvibile: **ci manca informazione**. Date solamente le variabili $\{c_0, \ldots, c_{i-1}\}$ e il prefisso X_i (del quale, in realtà, ci interessa solo l'elemento x_i), non c'è alcun modo per poter comprendere se l'elemento x_i possa essere accodato alle sottosequenze crescenti relative ai sottoproblemi di dimensione minore a i: non sappiamo con quale elemento termini ognuna di queste sottosequenze ma ne conosciamo solo la lunghezza.

Risulta necessario introdurre un problema ausiliario, nel quale introdurre l'informazione mancante necessaria.

2.3 Problema ausiliario `P'` di `PBR`

Il problema ausiliario `P'` è definito come segue:

data una sequenza X di m numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze crescenti di X <u>e che termina con x_m</u>

In altre parole, si richiede che la sottosequenza crescente più lunga termini con l'ultimo elemento della sequenza in input.

Come il problema ridotto `PBR`, anche il problema ausiliario contiene $m+1$ diversi sottoproblemi, ognuno associato ad una differente variabile. Il sottoproblema di dimensione (i) è definito come:

data una sequenza X di m numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze crescenti di X_i e che termina con x_i

ed è associato alla variabile c_i così definita:

c_i = *lunghezza di una tra le più lunghe sottosequenze crescenti di X_i e che termina con x_i*

Si noti quindi che per un qualunque sottoproblema di dimensione (s), solo imponendo che la soluzione c_s si riferisca ad una sottosequenza crescente di X_s che termina con x_s è possibile stabilire se un altro elemento x_u con $u > s$ possa essere accodato a tale sottosequenza (andando a verificare che $x_s < x_u$).

2.4 Equazioni di ricorrenza

É ora possibile risolvere il problema ausiliario `P'` andando a definire le **equazioni di ricorrenza**.

Le equazioni di ricorrenza si scrivono senza sapere il valore delle soluzioni dei sottoproblemi ma sapendo solamente che tali soluzioni esistono e si possono utilizzare

Un'equazione di ricorrenza è composta da:

- un **caso base** che definisce i casi più semplici che possono essere subito risolti senza ricorrere alle soluzioni dei sottoproblemi più piccoli
- un **passo ricorsivo** che definisce come risolvere i casi più complessi a partire dalle soluzioni dei sottoproblemi più piccoli (che si assume essere stati già risolti)

2.4.1 Caso base: (i) con $i \leq 1$

Il caso base si ha per un qualunque sottoproblema di dimensione (i) con $i = 0 \vee i = 1$, ossia quando il prefisso considerato è la sequenza vuota oppure è una sequenza composta da un singolo elemento. In entrambi i casi, è facile ottenere il valore della variabile c_i: la lunghezza di una tra le più lunghe sottosequenze crescenti della sequenza vuota è 0 mentre la lunghezza di una tra le più lunghe sottosequenze crescenti di una sequenza composta da un singolo elemento e che termina con quell'elemento è 1. Il caso base $i = 0$, comunque, risulta non necessario in quanto è possibile utilizzare, come caso base, solamente il caso $i = 1$. Per questa ragione, il caso base è scrivibile come:

$$c_i = 1 \quad \text{se } i = 1$$

i	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
x_i	ε	2	4	7	6	11	13	21	14	1
	0	1	2	3	3	4	5	6	6	1

Figura 3: Sottosequenza crescente: vettore di tutti i c_i ottenuti considerando la sequenza $X = \langle 2, 4, 7, 6, 11, 13, 21, 14, 1\rangle$. In rosso è evidenziata la soluzione del problema ridotto

2.4.2 Passo ricorsivo: (i) con $i > 1$

Il passo ricorsivo si ha per un qualunque sottoproblema di dimensione (i) con $i > 1$, ossia quando si considera un prefisso della sequenza X in input di almeno due elementi. I dati disponibili per calcolare c_i sono: l'input X ed in particolare l'elemento x_i e tutte le variabili $\{c_0, \ldots, c_{i-1}\}$. Si ricorda che i valori $c_1, c_2, \ldots, c_{i-1}$ rappresentano le lunghezze delle più lunghe sottosequenze crescenti di $X_1, X_2, \ldots, X_{i-1}$ che terminano, rispettivamente, con $x_1, x_2, \ldots, x_{i-1}$. Tra queste ci saranno alcune sottosequenze alle quali possiamo accodare x_i (in quanto maggiore dell'ultimo elemento) e altre alle quali l'elemento x_i non può essere accodato (in quanto minore dell'ultimo elemento). Se prendiamo la più lunga sottosequenza crescente alla quale possiamo attaccare x_i e accodiamo ad essa x_i, otteniamo la più lunga sottosequenza crescente di X_i che termina con x_i. La lunghezza di tale sottosequenza sarà uguale alla lunghezza della sottosequenza crescente alla quale abbiamo accodato x_i aumentata di 1. Il passo ricorsivo è quindi scrivibile come:

$$c_i = 1 + \max\{c_h \mid 1 \leq h < i \wedge x_h < x_i\}$$

Poiché può accadere che l'insieme $\{c_h \mid 1 \leq h < i \wedge x_h < x_i\}$ sia vuoto (il che corrisponde al fatto che l'elemento x_i è minore di tutti gli elementi precedenti e, quindi, non può essere accodato a nessuna più lunga sottosequenza crescente relativa a sottoproblemi di dimensione minore), assumiamo per definizione che $\max\{\emptyset\} = 0$, così che il corrispondente valore di c_i risulti uguale a 1.

2.5 Soluzione del problema

Una volta calcolati i valori $c_0, c_1, \ldots, c_m$ si hanno a disposizione tutte le lunghezza di una tra le più lunghe sottosequenze crescenti dei vari prefissi di X e che terminano con l'ultimo elemento del prefisso. La soluzione del problema ausiliario `P'` è c_m mentre quella del problema ridotto `PBR` è

$$\max\{c_i \mid 1 \leq i \leq m\}$$

Esempio Data $X = \langle 2, 4, 7, 6, 11, 13, 21, 14, 1\rangle$, tutti i valori c_i relativi sono contenuti nel vettore presentato in Figura 3.

2.6 Implementazione

L'algoritmo 2 presenta lo pseudocodice che calcola tutti i valori $c_0, c_1, \dots, c_m$ basandosi sulle equazioni di ricorrenza definite nella Sezione 2.4. Anziché utilizzare la ricorsione pura, tutti i valori sono calcolati con tecnica **bottom-up** in modo tale da risolvere ogni sottoproblema una volta sola. Questo permette di risolvere il problema in $\mathcal{O}(m^2)$ occupando $\Theta(m)$ spazio in memoria (ossia un vettore che contiene i vari c_i).

Algorithm 2 Lunghezza di una più lunga sottosequenza crescente di una sequenza X di lunghezza m.

```
procedure LGS(X)
    c[1] ← 1
    max ← c[1]
    for i ← 2 to m do
        temp ← 0
        for h ← 1 to i − 1 do
            if (x_h < x_i) ∧ (c[h] > temp) then
                temp ← c[h]
        c[i] ← 1 + temp
        if c[i] > max then
            max ← c[i]
    return max
```

3 Più lunga sottosequenza comune crescente (LGCS)

3.1 Definizione del problema

Il problema `PB` è definito come segue:

date due sequenze X e Y, rispettivamente di m e n numeri interi, si determini ***UNA*** *tra le più lunghe sottosequenze crescenti comuni a X e Y*

Ad esempio, date $X = \langle 2, 4, 7, 11, 21, 14, 1 \rangle$ e $Y = \langle 2, 7, 4, 23, 21, 14, 1, 8 \rangle$, una tra le più lunghe sottosequenze crescenti comuni è la sottosequenza $\langle 2, 4, 21 \rangle$.

Quello che si andrà a risolvere, comunque, è una versione ridotta del problema `PB`. La versione ridotta `PBR` è definita come segue:

date due sequenze X e Y, rispettivamente di m ed n numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze crescenti comuni a X e Y

Ad esempio, date $X = \langle 2, 4, 7, 11, 21, 14, 1 \rangle$ e $Y = \langle 2, 7, 4, 23, 21, 14, 1, 8 \rangle$, la lunghezza di una tra le più lunghe sottosequenze crescenti comuni è 3.

3.2 Sottoproblemi e variabili associate

Il problema `PBR` contiene diversi sottoproblemi ognuno dei quali non ha come input la coppia (X, Y) ma una coppia di prefissi di tali sequenze.

Il **sottoproblema di dimensione** (i, j) è definito come segue:

date due sequenze X e Y, rispettivamente di m ed n numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze crescenti comuni al prefisso X_i e al prefisso Y_j

Dato che $0 \leq i \leq m$ e $0 \leq j \leq n$, si ottengono $(m+1) \cdot (n+1)$ sottoproblemi (i e j possono valere 0 in quanto si deve considerare anche il caso in cui un prefisso sia la sequenza vuota). Ad ogni sottoproblema di `PBR` è associata una **variabile**.

Considerato il sottoproblema di dimensione (i, j), la variabile ad esso associata è $c_{i,j}$ ed è così definita:

$c_{i,j}$ = *lunghezza di una tra le più lunghe sottosequenze crescenti comuni a X_i e Y_j*

Per determinare la soluzione di un qualsiasi sottoproblema di dimensione (i, j), oltre all'input del problema, si utilizzeranno le soluzioni dei sottoproblemi di dimensione minore. Per esempio, dato il sottoproblema di dimensione $(2, 2)$, il valore della variabile $c_{2,2}$ si otterrà utilizzando le variabili $c_{0,0}$, $c_{0,1}$, $c_{1,0}$, $c_{1,1}$, $c_{1,2}$ e $c_{2,1}$. Si noti, però, che ognuna di queste variabili è da considerare come una ***black-box***: si può utilizzare ma non è possibile conoscerne il contenuto.

Considerato un qualunque sottoproblema di dimensione (i, j), assumendo di aver già risolto tutti i sottoproblemi di dimensione minore (ossia sapendo di poter utilizzare le variabili $\{c_{0,0}, \ldots, c_{i-1,j}, c_{i,j-1}\}$ senza però sapere effettivamente quanto valgono), come è possibile ottenere la soluzione $c_{i,j}$ andando a combinare le soluzioni dei sottoproblemi di dimensione minore a (i, j) ed, eventualmente, accodando gli elementi x_i e y_j estratti dall'input (X_i, Y_i)?

Purtroppo, però, il problema così definito non è risolvibile: **ci manca informazione**. Date solamente le variabili $\{c_{0,0}, \ldots, c_{i-1,j}, c_{i,j-1}\}$ e i prefissi X_i e Y_j (dei quali, in realtà, ci interessano solo gli elementi x_i e y_j), non c'è alcun modo per poter comprendere se gli elementi x_i e y_j, nel caso fossero uguali, possano essere accodati alle sottosequenze crescenti relative ai sottoproblemi di dimensione minore a (i, j): non sappiamo con quale elemento termini ognuna di queste sottosequenze ma ne conosciamo solo la lunghezza.

Risulta necessario introdurre un problema ausiliario, nel quale introdurre l'informazione mancante necessaria.

3.3 Problema ausiliario `P'` di `PBR`

Il problema ausiliario `P'` è definito come segue:

date due sequenze X e Y, rispettivamente di m ed n numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze crescenti comuni a X e a Y e che termina con x_m e y_n (se questi coincidono)

In altre parole, si richiede che la sottosequenza soluzione del problema termini con l'ultimo elemento delle sequenze X e Y nel caso siano lo stesso elemento (in caso contrario, la soluzione è la sequenza vuota).

Come il problema ridotto `PBR`, anche il problema ausiliario contiene $(m+1) \cdot (n+1)$ diversi sottoproblemi, ognuno associato ad una differente variabile. Il sottoproblema di dimensione (i, j) è definito come:

date due sequenze X e Y, rispettivamente di m ed n numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze crescenti comuni a X_i e a Y_j e che termina con x_i e y_j (se questi coincidono)

ed è associato alla variabile $c_{i,j}$ così definita:

$c_{i,j}$ = *lunghezza di una tra le più lunghe sottosequenze crescenti comuni a X_i e Y_j e che termina con x_i e y_j (se questi coincidono)*

Si noti quindi che per un qualunque sottoproblema di dimensione (s, t), solo imponendo che la soluzione $c_{s,t}$ si riferisca ad una sottosequenza crescente comune a X_s e Y_t che termina con $x_s = y_t$ è possibile stabilire se un altro elemento $x_u = y_v$ con $u > s$ e $v > t$ possa essere accodato a tale sottosequenza (andando a verificare che $x_s = y_t \leq x_u = y_v$).

3.4 Equazioni di ricorrenza

É ora possibile risolvere il problema ausiliario `P'` andando a definire le **equazioni di ricorrenza**.

Le equazioni di ricorrenza si scrivono senza sapere il valore delle soluzioni dei sottoproblemi ma sapendo solamente che tali soluzioni esistono e si possono utilizzare

Un'equazione di ricorrenza è composta da:

- un **caso base** che definisce i casi più semplici che possono essere subito risolti senza ricorrere alle soluzioni dei sottoproblemi più piccoli
- un **passo ricorsivo** che definisce come risolvere i casi più complessi a partire dalle soluzioni dei sottoproblemi più piccoli (che si assume essere stati già risolti)

3.4.1 Caso base

Certamente fanno parte del caso base tutte le coppie (i, j) con $i = 0 \vee j = 0$ ma il caso base si ha anche per un qualunque sottoproblema di dimensione (i, j) tale che $x_i \neq y_j$, ossia quando i due prefissi X_i e Y_j considerati terminano con due elementi diversi. Questo caso, infatti, è semplice da risolvere: per definizione di $c_{i,j}$ gli elementi x_i e y_j devono coincidere, in caso contrario la soluzione è 0 (ossia la lunghezza della sequenza vuota). Pertanto, si può evitare di considerare tutte le coppie (i, j) con $i = 0 \vee j = 0$: il numero di sottoproblemi che compone il problema ausiliario `P'` possono essere ridotti a $m \cdot n$.

Riassumendo, il caso base si ha per quei

sottoproblemi di dimensione (i, j) con $i > 0 \wedge j > 0$ tali che $x_i \neq y_j$

ed è scrivibile come

$$c_{i,j} = 0$$

3.4.2 Passo ricorsivo: (i, j) con $i > 0$ e $j > 0$ tali che $x_i = y_j$

Il passo ricorsivo si ha per un qualunque sottoproblema di dimensione (i, j) tale che $x_i = y_j$, ossia quando i due prefissi X_i e Y_j considerati terminano con lo stesso elemento. In questo caso, la lunghezza della più lunga sottosequenza crescente comune fra X_i e Y_j è uguale alla lunghezza della più lunga sottosequenza crescente comune calcolata per un sottoproblema di dimensione minore e che termina con un carattere minore di $x_i = y_j$ aumentata di uno (in quanto si accoda l'elemento comune $x_i = y_j$ alla sottosequenza crescente comune relativa al sottoproblema considerato). Il tutto può essere scritto come:

$$c_{i,j} = 1 + \max\{c_{h,k} \mid 1 \leq h < i, 1 \leq k < j, x_h < x_i\}$$

Poiché può accadere che l'insieme $\{c_{h,k} \mid 1 \leq h < i, 1 \leq k < j, x_h < x_i\}$ sia vuoto (il che corrisponde al fatto che l'elemento $x_i = y_j$ è minore di tutti gli elementi precedenti e, quindi, non può essere accodato a nessuna più lunga sottosequenza crescente relativa a sottoproblemi di dimensione minore), assumiamo per definizione che $\max\{\emptyset\} = 0$, così che il corrispondente valore di $c_{i,j}$ risulti uguale a 1. Inoltre, si noti che l'insieme $\{c_{h,k} \mid 1 \leq h < i, 1 \leq k < j, x_h < x_i\}$ corrisponde all'insieme vuoto anche se $i = 1 \vee j = 1$ in quanto i casi con $i = 0 \vee j = 0$ sono stati omessi dal caso base e, pertanto, non esistono sottoproblemi di dimensione minore a quella del sottoproblema considerato (ossia quello con $i = 1 \vee j = 1$).

3.5 Soluzione del problema

Una volta calcolati i valori $c_{1,1}, c_{2,1}, c_{1,2}, \ldots, c_{m,n}$ si hanno a disposizione tutte le lunghezze delle sottosequenze comuni massimali fra qualsiasi prefisso di X e qualsiasi prefisso di Y che non hanno elementi consecutivi dello stesso colore e che terminano con l'ultimo elemento di entrambi i prefissi. La soluzione al problema ausiliario `P'` è $c_{m,n}$ mentre quella del problema ridotto `PBR` è:

$$\max\{c_{i,j} \mid 1 \leq i \leq m \wedge 1 \leq j \leq n\}$$

Esempio Date $X = \langle 2, 4, 7, 11, 21, 14, 1\rangle$ e $Y = \langle 2, 7, 4, 23, 21, 14, 1, 8\rangle$, tutti i valori $c_{i,j}$ relativi sono contenuti nella matrice presentata in Figura 4.

3.6 Implementazione

L'algoritmo 3 presenta lo pseudocodice che calcola tutti i valori $c_{0,0}, c_{0,1}, \ldots, c_{m,n}$ basandosi sulle equazioni di ricorrenza definite nella Sezione 3.4. Anziché utilizzare la ricorsione pura, tutti i valori sono calcolati con tecnica **bottom-up** in modo tale da risolvere ogni sottoproblema una volta sola. Questo permette di risolvere il problema in $\mathcal{O}(m^2 \cdot n^2)$ occupando $\Theta(m \cdot n)$ spazio in memoria (ossia una matrice che contiene i vari $c_{i,j}$).

		1	2	3	4	5	6	7	8	**j**
		2	7	4	23	21	14	1	8	**y_j**
1	2	1	0	0	0	0	0	0	0	
2	4	0	0	2	0	0	0	0	0	
3	7	0	2	0	0	0	0	0	0	
4	11	0	0	0	0	0	0	0	0	
5	21	0	0	0	0	3	0	0	0	
6	14	0	0	0	0	0	3	0	0	
7	1	0	0	0	0	0	0	1	0	
i	**x_i**									

Figura 4: Sottosequenza crescente comune: matrice di tutti i $c_{i,j}$ ottenuti considerando le sequenze $X = \langle 2, 4, 7, 11, 21, 14, 1 \rangle$ e $Y = \langle 2, 7, 4, 23, 21, 14, 1, 8 \rangle$. In rosso è evidenziata la soluzione del problema ridotto

Algorithm 3 Lunghezza di una più lunga sottosequenza crescente comune fra due sequenze X e Y di lunghezza m e n rispettivamente.

```
procedure LGCS(X, Y)
    max ← 0
    for i ← 1 to m do
        for j ← 1 to n do
            if x_i ≠ y_j then
                c[i, j] ← 0
            else
                temp ← 0
                for h ← 1 to i − 1 do
                    for k ← 1 to j − 1 do
                        if (x_h < x_i) ∧ (c[h, k] > temp) then
                            temp ← c[h, k]
                c[i, j] ← 1 + temp
            if c[i, j] > max then
                max ← c[i, j]
    return max
```

4 Più lunga sottosequenza con colori alternanti (LAS)

Premessa Sia C un insieme di colori. Gli elementi della sequenza considerata in questo esercizio sono numeri naturali colorati. La funzione di colorazione è definita nel seguente modo: $\phi : \mathbb{N} \rightarrow C$. Per esempio, data la funzione:

$$\phi(x) = \begin{cases} \text{rosso} & x < 5 \\ \text{blu} & 5 \leq x \leq 10 \\ \text{verde} & x > 10 \end{cases}$$

la sequenza $X = \langle 2, 4, 7, 6, 11, 3, 21, 14, 1 \rangle$ sarà colorata nel seguente modo: $\langle r, r, b, b, g, r, g, g, r \rangle$.

4.1 Definizione del problema

Il problema `PB` è definito come segue:

data una sequenza X di m numeri interi, si determini ***UNA*** *tra le più lunghe sottosequenze di X che non abbia elementi consecutivi dello stesso colore*

Ad esempio, data $X = \langle 2, 4, 7, 6, 11, 3, 21, 14, 12 \rangle$ colorata utilizzando la funzione ϕ definita sopra, una tra le più lunghe sottosequenze di X che non ha elementi consecutivi dello stesso colore è la sottosequenza $\langle 2, 7, 11, 3, 21 \rangle$.

Quello che si andrà a risolvere, comunque, è una versione ridotta del problema `PB`. La versione ridotta `PBR` è definita come segue:

data una sequenza X di m numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze di X che non abbia elementi consecutivi dello stesso colore

Ad esempio, data $X = \langle 2, 4, 7, 6, 11, 3, 21, 14, 12 \rangle$, la lunghezza di una tra le più lunghe sottosequenze di X che non ha elementi consecutivi dello stesso colore è è 5.

4.2 Sottoproblemi e variabili associate

Il problema `PBR` contiene diversi sottoproblemi ognuno dei quali non ha come input la sequenza X ma un suo prefisso.

Il **sottoproblema di dimensione** (i) è definito come segue:

data una sequenza X di m numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze di X_i che non abbia elementi consecutivi dello stesso colore

Dato che $0 \leq i \leq m$, si ottengono $(m + 1)$ sottoproblemi (i può valere 0 in quanto si deve considerare anche il caso in cui il prefisso sia la sequenza vuota). Ad ogni sottoproblema di `PBR` è associata una **variabile**.

Considerato il sottoproblema di dimensione (i), la variabile ad esso associata è c_i ed è così definita:

c_i = *lunghezza di una tra le più lunghe sottosequenze di X_i che non ha elementi consecutivi dello stesso colore*

Per determinare la soluzione di un qualsiasi sottoproblema di dimensione (i), oltre all'input del problema, si utilizzeranno le soluzioni dei sottoproblemi di dimensione minore. Per esempio, dato il sottoproblema di dimensione (3), il valore della variabile c_3 si otterrà utilizzando le variabili c_0, c_1 e c_2. Si noti, però, che ognuna di queste variabili è da considerare come una ***black-box***: si può utilizzare ma non è possibile conoscerne il contenuto.

Considerato un qualunque sottoproblema di dimensione (i), assumendo di aver già risolto tutti i sottoproblemi di dimensione minore (ossia sapendo di poter utilizzare le variabili $\{c_0, \ldots, c_{i-1}\}$ senza però sapere effettivamente quanto valgono), come è possibile ottenere la soluzione c_i andando a combinare le soluzioni dei sottoproblemi di dimensione minore a (i) ed, eventualmente, accodando l'elemento x_i estratto dall'input X_i?

Purtroppo, però, il problema così definito non è risolvibile: **ci manca informazione**. Date solamente le variabili $\{c_0, \ldots, c_{i-1}\}$ e il prefisso X_i (del quale, in realtà, ci interessa solo l'elemento x_i), non c'è alcun modo per poter comprendere se l'elemento x_i possa essere accodato alle sottosequenze senza elementi consecutivi dello stesso colore relative ai sottoproblemi di dimensione minore a i: non sappiamo con quale elemento (e quindi con quale colore) termini ognuna di queste sottosequenze ma ne conosciamo solo la lunghezza.

Risulta necessario introdurre un problema ausiliario, nel quale introdurre l'informazione mancante necessaria.

4.3 Problema ausiliario `P'` di `PBR`

Il problema ausiliario `P'` è definito come segue:

data una sequenza X di m numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze di X che non abbia elementi consecutivi dello stesso colore <u>e che termini con x_m</u>

In altre parole, si richiede che la sottosequenza più lunga senza elementi consecutivi dello stesso colore termini con l'ultimo elemento della sequenza in input.

Come il problema ridotto `PBR`, anche il problema ausiliario contiene $m+1$ diversi sottoproblemi, ognuno associato ad una differente variabile. Il sottoproblema di dimensione (i) è definito come:

data una sequenza X di m numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze di X_i che non abbia elementi consecutivi dello stesso colore e che termini con x_i

ed è associato alla variabile c_i così definita:

c_i = *lunghezza di una tra le più lunghe sottosequenze di X_i che non ha elementi consecutivi dello stesso colore e che termina con x_i*

Si noti quindi che per un qualunque sottoproblema di dimensione (s), solo imponendo che la soluzione c_s si riferisca ad una sottosequenza di X_s che non ha elementi consecutivi dello stesso colore e che termina con x_s è possibile stabilire se un altro elemento x_u con $u > s$ possa essere accodato a tale sottosequenza (andando a verificare che $\phi(x_s) \neq \phi(x_u)$).

4.4 Equazioni di ricorrenza

É ora possibile risolvere il problema ausiliario `P'` andando a definire le **equazioni di ricorrenza**.

Le equazioni di ricorrenza si scrivono senza sapere il valore delle soluzioni dei sottoproblemi ma sapendo solamente che tali soluzioni esistono e si possono utilizzare

Un'equazione di ricorrenza è composta da:

- un **caso base** che definisce i casi più semplici che possono essere subito risolti senza ricorrere alle soluzioni dei sottoproblemi più piccoli
- un **passo ricorsivo** che definisce come risolvere i casi più complessi a partire dalle soluzioni dei sottoproblemi più piccoli (che si assume essere stati già risolti)

4.4.1 Caso base: (i) con $i \leq 1$

Il caso base si ha per un qualunque sottoproblema di dimensione (i) con $i = 0 \lor i = 1$, ossia quando il prefisso considerato è la sequenza vuota oppure è una sequenza composta da un singolo elemento. In entrambi i casi, è facile ottenere il valore della variabile c_i: la lunghezza di una tra le più lunghe sottosequenze della sequenza vuota che non ha elementi consecutivi dello stesso colore è 0 (la lunghezza della sequenza vuota) mentre la lunghezza di una tra le più lunghe sottosequenze di una sequenza di un singolo elemento che non ha elementi consecutivi dello stesso colore e che termina con quell'elemento è 1. Il caso base $i = 0$, comunque, risulta non necessario in quanto è possibile utilizzare, come caso base, solamente il caso $i = 1$. Per questa ragione, il caso base è scrivibile come:

$$c_i = 1 \quad \text{se } i = 1$$

4.4.2 Passo ricorsivo: (i) con $i > 1$

Il passo ricorsivo si ha per un qualunque sottoproblema di dimensione (i) con $i > 1$, ossia quando si considera un prefisso della sequenza X in input di almeno due elementi. I dati disponibili per calcolare c_i sono: l'input X ed in particolare l'elemento x_i e tutte le variabili $\{c_0, \ldots, c_{i-1}\}$. Si ricorda che i valori $c_1, c_2, \ldots, c_{i-1}$ rappresentano le lunghezze delle più lunghe sottosequenze di $X_1, X_2, \ldots, X_{i-1}$ che non contengono elementi consecutivi dello stesso colore e che terminano, rispettivamente, con $x_1, x_2, \ldots, x_{i-1}$. Tra queste ci saranno alcune sottosequenze alle quali possiamo accodare x_i (in quanto colorato in modo diverso dell'ultimo elemento) e altre alle quali l'elemento x_i non può essere accodato (in quanto colorato allo stesso modo dell'ultimo elemento). Se prendiamo la più lunga sottosequenza alla quale possiamo attaccare x_i e accodiamo

i	0	1	2	3	4	5	6	7	8	9
$\mathbf{x_i}$	ε	2	4	7	6	11	3	21	14	12
$\boldsymbol{\varphi(x_i)}$	-	r	r	b	b	g	r	g	g	g
	0	1	1	2	2	3	4	5	5	5

Figura 5: Sottosequenza con colori alternanti: vettore di tutti i c_i ottenuti considerando la sequenza $X = \langle 2, 4, 7, 6, 11, 3, 21, 14, 12 \rangle$. In rosso è evidenziata la soluzione del problema ridotto

ad essa x_i, otteniamo la più lunga sottosequenza di X_i che non contiene elementi consecutivi dello stesso colore e che termina con x_i. La lunghezza di tale sottosequenza, pertanto, sarà uguale alla lunghezza della sottosequenza alla quale abbiamo accodato x_i aumentata di 1. Il passo ricorsivo è quindi scrivibile come:

$$c_i = 1 + \max\{c_h \mid 1 \leq h < i \wedge \phi(x_h) \neq \phi(x_i)\}$$

Poiché può accadere che l'insieme $\{c_h \mid 1 \leq h < i \wedge \phi(x_h) \neq \phi(x_i)\}$ sia vuoto (il che corrisponde al fatto che l'elemento x_i è colorato allo stesso modo di tutti gli elementi precedenti e, quindi, non può essere accodato a nessuna più lunga sottosequenza alternante relativa a sottoproblemi di dimensione minore), assumiamo per definizione che $\max\{\emptyset\} = 0$, così che il corrispondente valore di c_i risulti uguale a 1.

4.5 Soluzione del problema

Una volta calcolati i valori $c_0, c_1, \ldots, c_m$ si hanno a disposizione le lunghezze di una tra le più lunghe sottosequenze dei vari prefissi della sequenza in input che non contengono elementi consecutivi dello stesso colore e che terminano con l'ultimo elemento del prefisso. La soluzione al problema ausiliario `P'` è c_m mentre quella del problema ridotto `PBR` è

$$\max\{c_i \mid 1 \leq i \leq m\}$$

Esempio Data $X = \langle 2, 4, 7, 6, 11, 3, 21, 14, 12 \rangle$, tutti i valori c_i relativi sono contenuti nel vettore presentato in Figura 5.

4.6 Implementazione

L'algoritmo 2 presenta lo pseudocodice che calcola tutti i valori $c_0, c_1, \ldots, c_m$ basandosi sulle equazioni di ricorrenza definite nella Sezione 2.4. Anziché utilizzare la ricorsione pura, tutti i valori sono calcolati con tecnica **bottom-up** in modo tale da risolvere ogni sottoproblema una volta sola. Questo permette di risolvere il problema in $\mathcal{O}(m^2)$ occupando $\Theta(m)$ spazio in memoria (ossia un vettore che contiene i vari c_i).

Algorithm 4 Lunghezza di una più lunga sottosequenza di una sequenza X lunga m che non presenta elementi consecutivi dello stesso colore.

```
procedure LAS(X)
    c[1] ← 1
    max ← c[1]
    for i ← 2 to m do
        temp ← 0
        for h ← 1 to i − 1 do
            if (φ(x_h) ≠ φ(x_i)) ∧ (c[h] > temp) then
                temp ← c[h]
        c[i] ← 1 + temp
        if c[i] > max then
            max ← c[i]
    return max
```

5 Più lunga sottosequenza comune con colore alternante (LACS)

Premessa Sia C un insieme di colori. Gli elementi della sequenza considerata in questo esercizio sono numeri naturali colorati. La funzione di colorazione è definita nel seguente modo: $\phi : \mathbb{N} \to C$. Per esempio, data la funzione:

$$\phi(x) = \begin{cases} \text{rosso} & x < 5 \\ \text{blu} & 5 \leq x \leq 10 \\ \text{verde} & x > 10 \end{cases}$$

la sequenza $X = \langle 2, 4, 7, 6, 11, 3, 21, 14, 1 \rangle$ sarà colorata nel seguente modo: $\langle r, r, b, b, g, r, g, g, r \rangle$.

5.1 Definizione del problema

Il problema `PB` è definito come segue:

date due sequenze X e Y, rispettivamente di m e n numeri interi, si determini ***UNA*** *tra le più lunghe sottosequenze comuni a X e Y che non abbia elementi consecutivi dello stesso colore*

Ad esempio, date $X = \langle 2, 4, 7, 11, 21, 14, 1 \rangle$ e $Y = \langle 2, 7, 4, 23, 21, 14, 1, 8 \rangle$, una tra le più lunghe sottosequenze comuni senza elementi consecutivi dello stesso colore è la sottosequenza $\langle 2, 7, 21, 1 \rangle$.

Quello che si andrà a risolvere, comunque, è una versione ridotta del problema `PB`. La versione ridotta `PBR` è definita come segue:

date due sequenze X e Y, rispettivamente di m ed n numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze comuni a X e Y che non abbia elementi consecutivi dello stesso colore

Ad esempio, date $X = \langle 2, 4, 7, 11, 21, 14, 1 \rangle$ e $Y = \langle 2, 7, 4, 23, 21, 14, 1, 8 \rangle$, la lunghezza di una tra le più lunghe sottosequenze comuni senza elementi consecutivi dello stesso colore è 4.

5.2 Sottoproblemi e variabili associate

Il problema `PBR` contiene diversi sottoproblemi ognuno dei quali non ha come input la coppia (X, Y) ma una coppia di prefissi di tali sequenze.

Il **sottoproblema di dimensione** (i, j) è definito come segue:

date due sequenze X e Y, rispettivamente di m ed n numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze comuni a X_i e Y_j che non abbia elementi consecutivi dello stesso colore

Dato che $0 \leq i \leq m$ e $0 \leq j \leq n$, si ottengono $(m+1)\cdot(n+1)$ sottoproblemi (i e j possono valere 0 in quanto si deve considerare anche il caso in cui un prefisso sia la sequenza vuota). Ad ogni sottoproblema di `PBR` è associata una **variabile**.

Considerato il sottoproblema di dimensione (i, j), la variabile ad esso associata è $c_{i,j}$ ed è così definita:

$c_{i,j}$ = *lunghezza di una tra le più lunghe sottosequenze comuni a X_i e Y_j che non ha elementi consecutivi dello stesso colore*

Per determinare la soluzione di un qualsiasi sottoproblema di dimensione (i, j), oltre all'input del problema, si utilizzeranno le soluzioni dei sottoproblemi di dimensione minore. Per esempio, dato il sottoproblema di dimensione $(2, 2)$, il valore della variabile $c_{2,2}$ si otterrà utilizzando le variabili $c_{0,0}$, $c_{0,1}$, $c_{1,0}$, $c_{1,1}$, $c_{1,2}$ e $c_{2,1}$. Si noti, però, che ognuna di queste variabili è da considerare come una ***black-box***: si può utilizzare ma non è possibile conoscerne il contenuto.

Considerato un qualunque sottoproblema di dimensione (i, j), assumendo di aver già risolto tutti i sottoproblemi di dimensione minore (ossia sapendo di poter utilizzare le variabili $\{c_{0,0}, \ldots, c_{i-1,j}, c_{i,j-1}\}$ senza però sapere effettivamente quanto valgono), come è possibile ottenere la soluzione $c_{i,j}$ andando a combinare le soluzioni dei sottoproblemi di dimensione minore a (i, j) ed, eventualmente, accodando gli elementi x_i e y_j estratti dall'input (X_i, Y_i)?

Purtroppo, però, il problema così definito non è risolvibile: **ci manca informazione**. Date solamente le variabili $\{c_{0,0}, \ldots, c_{i-1,j}, c_{i,j-1}\}$ e i prefissi X_i e Y_j (dei quali, in realtà, ci interessano solo gli elementi x_i e y_j), non c'è alcun modo per poter comprendere se gli elementi x_i e y_j, nel caso fossero uguali, possano essere accodati alle sottosequenze comuni alternanti relative ai sottoproblemi di dimensione minore a (i, j): non sappiamo con quale elemento (e quindi con quale colore) termini ognuna di queste sottosequenze ma ne conosciamo solo la lunghezza.

Risulta necessario introdurre un problema ausiliario, nel quale introdurre l'informazione mancante necessaria.

5.3 Problema ausiliario `P'` di `PBR`

Il problema ausiliario `P'` è definito come segue:

date due sequenze X e Y, rispettivamente di m ed n numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze comuni a X e Y

che non abbia elementi consecutivi dello stesso colore
e che termini con x_m e y_n (se questi coincidono)

In altre parole, si richiede che la sottosequenza soluzione del problema termini con l'ultimo elemento delle sequenze X e Y nel caso siano lo stesso elemento (in caso contrario, la soluzione è la sequenza vuota).

Come il problema ridotto `PBR`, anche il problema ausiliario contiene $(m + 1) \cdot (n + 1)$ diversi sottoproblemi, ognuno associato ad una differente variabile. Il sottoproblema di dimensione (i, j) è definito come:

date due sequenze X e Y, rispettivamente di m ed n numeri interi, si determini la lunghezza di una tra le più lunghe sottosequenze comuni a X_i e Y_j che non abbia elementi consecutivi dello stesso colore e che termini con x_i e y_j (se questi coincidono)

ed è associato alla variabile $c_{i,j}$ così definita:

$c_{i,j}$ = *lunghezza di una tra le più lunghe sottosequenze comuni a X_i e Y_j che non ha elementi consecutivi dello stesso colore e che termina con x_i e y_j (se questi coincidono)*

Si noti quindi che per un qualunque sottoproblema di dimensione (s, t), solo imponendo che la soluzione $c_{s,t}$ si riferisca ad una sottosequenza comune a X_s e Y_t che non presenta elementi consecutivi dello stesso colore e che termina con $x_s = y_t$ è possibile stabilire se un altro elemento $x_u = y_v$ con $u > s$ e $v > t$ possa essere accodato a tale sottosequenza (andando a verificare che $\phi(x_s) = \phi(y_t) \neq \phi(x_u) = \phi(y_v)$).

5.4 Equazioni di ricorrenza

É ora possibile risolvere il problema ausiliario `P'` andando a definire le **equazioni di ricorrenza**.

Le equazioni di ricorrenza si scrivono senza sapere il valore delle soluzioni dei sottoproblemi ma sapendo solamente che tali soluzioni esistono e si possono utilizzare

Un'equazione di ricorrenza è composta da:

- un **caso base** che definisce i casi più semplici che possono essere subito risolti senza ricorrere alle soluzioni dei sottoproblemi più piccoli

- un **passo ricorsivo** che definisce come risolvere i casi più complessi a partire dalle soluzioni dei sottoproblemi più piccoli (che si assume essere stati già risolti)

5.4.1 Caso base

Certamente fanno parte del caso base tutte le coppie (i, j) con $i = 0 \vee j = 0$ ma il caso base si ha anche per un qualunque sottoproblema di dimensione (i, j) tale che $x_i \neq y_j$, ossia quando i due prefissi X_i e Y_j considerati terminano con due elementi diversi. Questo caso, infatti, è semplice da risolvere: per definizione di

$c_{i,j}$ gli elementi x_i e y_j devono coincidere, in caso contrario la soluzione è 0 (ossia la lunghezza della sequenza vuota). Pertanto, si può evitare di considerare tutte le coppie (i, j) con $i = 0 \vee j = 0$: il numero di sottoproblemi che compone il problema ausiliario `P'` possono essere ridotti a $m \cdot n$.

Riassumendo, il caso base si ha per quei

sottoproblemi di dimensione (i, j) con $i > 0 \wedge j > 0$ tali che $x_i \neq y_j$

ed è scrivibile come

$$c_{i,j} = 0$$

5.4.2 Passo ricorsivo: (i, j) con $i > 0$ e $j > 0$ tali che $x_i = y_j$

Il passo ricorsivo si ha per un qualunque sottoproblema di dimensione (i, j) tale che $x_i = y_j$, ossia quando i due prefissi X_i e Y_j considerati terminano con lo stesso elemento. In questo caso, la lunghezza della più lunga sottosequenza con colori alternanti comune fra X_i e Y_j è uguale alla lunghezza della più lunga sottosequenza comune con colori alternanti calcolata per un sottoproblema di dimensione minore e che termina con un carattere con colore diverso da quello di $x_i = y_j$ aumentata di uno (in quanto si accoda l'elemento comune x_i alla sottosequenza comune relativa al sottoproblema considerato). Il tutto può essere scritto come:

$$c_{i,j} = 1 + \max\{c_{h,k} \mid 1 \leq h < i, 1 \leq k < j, \phi(x_i) \neq \phi(x_h)\}$$

Poiché può accadere che l'insieme $\{c_{h,k} \mid 1 \leq h < i, 1 \leq k < j, \phi(x_i) \neq \phi(x_h)\}$ sia vuoto (il che corrisponde al fatto che l'elemento $x_i = y_j$ è colorato allo stesso modo di tutti gli elementi precedenti e, quindi, non può essere accodato a nessuna più lunga sottosequenza alternante relativa a sottoproblemi di dimensione minore), assumiamo per definizione che $\max\{\emptyset\} = 0$, così che il corrispondente valore di $c_{i,j}$ risulti uguale a 1. Inoltre, si noti che l'insieme $\{c_{h,k} \mid 1 \leq h < i, 1 \leq k < j, x_h < x_i\}$ corrisponde all'insieme vuoto anche se $i = 1 \vee j = 1$ in quanto i casi con $i = 0 \vee j = 0$ sono stati omessi dal caso base e, pertanto, non esistono sottoproblemi di dimensione minore a quella del sottoproblema considerato (ossia quello con $i = 1 \vee j = 1$).

5.5 Soluzione del problema

Una volta calcolati i valori $c_{1,1}, c_{2,1}, c_{1,2}, \ldots, c_{m,n}$ si hanno a disposizione tutte le lunghezze delle sottosequenze comuni massimali fra qualsiasi prefisso di X e qualsiasi prefisso di Y che non hanno elementi consecutivi dello stesso colore e che terminano con l'ultimo elemento di entrambi i prefissi. La soluzione al problema ausiliario `P'` è $c_{m,n}$ mentre quella del problema ridotto `PBR` è:

$$\max\{c_{i,j} \mid 1 \leq i \leq m \wedge 1 \leq j \leq n\}$$

Esempio Date $X = \langle 2, 4, 7, 11, 21, 14, 1\rangle$ e $Y = \langle 2, 7, 4, 23, 21, 14, 1, 8\rangle$, tutti i valori $c_{i,j}$ relativi sono contenuti nella matrice presentata in Figura 6.

		j	1	2	3	4	5	6	7	8
		y_j	2	7	4	23	21	14	1	8
		$\varphi(y_j)$	r	b	r	g	g	g	r	b
1	2	r	1	0	0	0	0	0	0	0
2	4	r	0	0	1	0	0	0	0	0
3	7	b	0	2	0	0	0	0	0	0
4	11	g	0	0	0	0	0	0	0	0
5	21	g	0	0	0	0	3	0	0	0
6	14	g	0	0	0	0	0	3	0	0
7	1	r	0	0	0	0	0	0	4	0
i	x_i	$\varphi(x_i)$								

Figura 6: Sottosequenza comune con colori alternanti: matrice di tutti i $c_{i,j}$ ottenuti considerando le sequenze $X = \langle 2, 4, 7, 11, 21, 14, 1 \rangle$ e $Y = \langle 2, 7, 4, 23, 21, 14, 1, 8 \rangle$. In rosso è evidenziata la soluzione del problema ridotto

5.6 Implementazione

L'algoritmo 5 presenta lo pseudocodice che calcola tutti i valori $c_{0,0}, c_{0,1}, \ldots, c_{m,n}$ basandosi sulle equazioni di ricorrenza definite nella Sezione 5.4. Anziché utilizzare la ricorsione pura, tutti i valori sono calcolati con tecnica **bottom-up** in modo tale da risolvere ogni sottoproblema una volta sola. Questo permette di risolvere il problema in $\mathcal{O}(m^2 \cdot n^2)$ occupando $\Theta(m \cdot n)$ spazio in memoria (ossia una matrice che contiene i vari $c_{i,j}$).

6 Più lunga sottosequenza comune di lunghezza almeno L (LCS≥L)

6.1 Definizione del problema

Il problema `PB` è definito come segue:

date due sequenze X e Y, rispettivamente di m e n numeri interi, e un naturale L, stabilire se la lunghezza di una qualunque tra le più lunghe sottosequenze comuni a X e Y è maggiore o uguale a L

Ad esempio, date $X = \langle 2, 4, 7, 11, 14, 21, 8 \rangle$, $Y = \langle 2, 7, 4, 23, 11, 14, 21, 1 \rangle$ e $L = 4$, la soluzione al problema deve essere *true* in quanto tutte le più lunghe sottosequenze comuni hanno lunghezza 5 (che è maggiore di 4).

Per risolvere il problema `PB` non è necessario introdurre alcuna versione ridotta.

Algorithm 5 Lunghezza di una più lunga sottosequenza con colori alternanti comune fra due sequenze X e Y di lunghezza m e n rispettivamente.

```
procedure LACS(X, Y)
    max ← 0
    for i ← 1 to m do
        for j ← 1 to n do
            if x_i ≠ y_j then
                c[i, j] ← 0
            else
                temp ← 0
                for h ← 1 to i − 1 do
                    for k ← 1 to j − 1 do
                        if (φ(x_h) ≠ φ(x_i)) ∧ (c[h, k] > temp) then
                            temp ← c[h, k]
                c[i, j] ← 1 + temp
            if c[i, j] > max then
                max ← c[i, j]
    return max
```

6.2 Sottoproblemi e variabili associate

Il problema PB contiene diversi sottoproblemi ognuno dei quali non ha come input la tripla (X, Y, L) ma una tripla composta dai prefissi delle sequenze X e Y e da un certo naturale $0 \leq l \leq L$.

Il **sottoproblema di dimensione** (i, j, l) è definito come segue:

date due sequenze X e Y, rispettivamente di m ed n numeri interi, e un naturale l, stabilire se la lunghezza di una qualunque tra le più lunghe sottosequenze comuni a X_i e Y_j è maggiore o uguale a l

Dato che $0 \leq i \leq m$, $0 \leq j \leq n$ e $0 \leq l \leq L$, si ottengono $(m+1) \cdot (n+1) \cdot (L+1)$ sottoproblemi (i e j possono valere 0 in quanto si deve considerare anche il caso in cui un prefisso sia la sequenza vuota). Ad ogni sottoproblema di PBR è associata una **variabile**.

Considerato il sottoproblema di dimensione (i, j, l), la variabile ad esso associata è $c_{i,j,l}$ ed è così definita:

$c_{i,j,l} =$ ***true*** *se e solo se la lunghezza di una qualunque tra le più lunghe sottosequenze comuni a X_i e Y_j è maggiore o uguale a l,* ***false*** *altrimenti*

Per determinare la soluzione di un qualsiasi sottoproblema di dimensione (i, j, l), oltre all'input del problema, si utilizzeranno le soluzioni dei sottoproblemi di dimensione minore. Per esempio, dato il sottoproblema di dimensione $(1, 1, 1)$, il valore della variabile $c_{1,1,1}$ si otterrà utilizzando le variabili $c_{0,0,0}$, $c_{0,0,1}$, $c_{1,0,0}$, $c_{1,0,1}$, $c_{0,1,0}$ e $c_{0,1,1}$. Si noti, però, che ognuna di queste variabili è da considerare come una ***black-box***: si può utilizzare ma non è possibile conoscerne il contenuto.

Considerato un qualunque sottoproblema di dimensione (i, j, l), assumendo di aver già risolto tutti i sottoproblemi di dimensione minore (ossia sapendo di poter utilizzare tutte le variabili $c_{i',j',l'}$ relative ai sottoproblemi di dimensione minore senza però sapere effettivamente quanto valgono), come è possibile ottenere la soluzione $c_{i,j,l}$ andando a combinare le soluzioni dei sottoproblemi di dimensione minore a (i, j, l) ed, eventualmente, accodando gli elementi x_i e y_j estratti dall'input (X_i, Y_i) e utilizzando il valore l?

6.3 Equazioni di ricorrenza

Per rispondere a questa domanda, si devono andare a definire le **equazioni di ricorrenza**.

Le equazioni di ricorrenza si scrivono senza sapere il valore delle soluzioni dei sottoproblemi ma sapendo solamente che tali soluzioni esistono e si possono utilizzare

Un'equazione di ricorrenza è composta da:

- un **caso base** che definisce i casi più semplici che possono essere subito risolti senza ricorrere alle soluzioni dei sottoproblemi più piccoli
- un **passo ricorsivo** che definisce come risolvere i casi più complessi a partire dalle soluzioni dei sottoproblemi più piccoli (che si assume essere stati già risolti)

6.3.1 Caso base

Il caso base si ha in due situazioni differenti:

1. per un qualunque sottoproblema di dimensione (i, j, l) con
$$l = 0$$
2. per un qualunque sottoproblema di dimensione (i, j, l) con
$$l > 0 \wedge (i = 0 \vee j = 0)$$

Per entrambi questi casi, infatti, è facile stabilire il valore di $c_{i,j,l}$.

Caso base 1: (i, j, l) **con** $l = 0$ Questo caso base si verifica quando si richiede che una qualunque tra le più lunghe sottosequenze comuni a X_i e Y_j abbia lunghezza maggiore o uguale a 0. È facile ottenere il valore della variabile $c_{i,j,0}$ in quanto vale sempre *true*: c'è sempre una sottosequenza comune di lunghezza maggiore o uguale a 0. Infatti, nel caso migliore si ha una sottosequenza comune di lunghezza maggiore a 0 mentre nel caso peggiore, si ha una sottosequenza comunque di lunghezza 0 (ossia la lunghezza di una sottosequenza comune fra la sequenza vuota e una qualsiasi altra sequenza). Questo caso base può essere scritto nel seguente modo:

$$c_{i,j,l} = true \quad se \;\; l = 0$$

Caso base 2: (i, j, l) **con** $l > 0 \wedge (i = 0 \vee j = 0)$ Questo caso base si verifica quando uno (o entrambi) i prefissi considerati sono uguali alla stringa vuota e si richiede che una qualunque tra le più lunghe sottosequenze comuni a X_i e Y_j abbia lunghezza maggiore a 0. Questo è impossibile in quanto la lunghezza di una qualunque tra le più lunghe sottosequenze comuni fra la sequenza vuota e una qualsiasi altra sequenza è pari a 0 (ossia la la lunghezza della sequenza vuota). Questo caso base può essere scritto nel seguente modo:

$$c_{i,j,l} = false \quad se \;\; l > 0 \wedge (i = 0 \vee j = 0)$$

Riassumendo, il caso base può essere scritto nel seguente modo:

$$c_{i,j,l} = \begin{cases} true & se \;\; l = 0 \\ false & se \;\; l > 0 \wedge (i = 0 \vee j = 0) \end{cases}$$

6.3.2 Passo ricorsivo

Il passo ricorsivo si ha nelle seguenti situazioni:

1. per un qualunque sottoproblema di dimensione (i, j, l) con

$$i > 0 \wedge j > 0 \wedge l > 0 \wedge x_i = y_j$$

2. per un qualunque sottoproblema di dimensione (i, j, l) con

$$i > 0 \wedge j > 0 \wedge l > 0 \wedge x_i \neq y_j$$

Passo ricorsivo 1: (i, j, l) **con** $i > 0 \wedge j > 0 \wedge l > 0$ ***tale che*** $x_i = y_j$ Questo passo ricorsivo si verifica quando si vanno a considerare due prefissi $X_i = \langle x_0, x_1, ..., x_i \rangle$ e $Y_j = \langle y_0, y_1, ..., y_j \rangle$ entrambi diversi dalla sequenza vuota che terminano con lo stesso elemento.

Se i due prefissi terminano con lo stesso elemento ($x_i = y_j$) allora la lunghezza di una qualunque tra le più lunghe sottosequenze comuni a X_i e Y_j è maggiore o uguale a l se e solo se la lunghezza di una qualunque tra le più lunghe sottosequenze comuni a X_{i-1} e Y_{j-1} risulta essere maggiore o uguale $l-1$ (si considera la soluzione $c_{i-1,j-1,l-1}$ relativa al sottoproblema di dimensione $(i-1, j-1, l-1)$ con input il prefisso X_{i-1}, il prefisso Y_{j-1} e la lunghezza $l-1$). Infatti, solo se esiste una più lunga sottosequenze comune a X_{i-1} e Y_{j-1} di lunghezza maggiore o uguale a $l-1$ allora è possibile concatenare ad essa l'elemento comune $x_i = y_j$ ed ottenere una sottosequenza comune a X_i e Y_j di lunghezza maggiore o uguale a l.

Riassumendo, questo passo ricorsivo può essere scritto come:

$$c_{i,j,l} = c_{i-1,j-1,l-1} \qquad se \; i > 0 \wedge j > 0 \wedge l > 0 \wedge x_i = y_j$$

Passo ricorsivo 2: (i, j, l) **con** $i > 0 \wedge j > 0 \wedge l > 0$ ***tale che*** $x_i \neq y_j$
Questo passo ricorsivo si verifica quando si vanno a considerare due prefissi $X_i = \langle x_0, x_1, ..., x_i \rangle$ e $Y_j = \langle y_0, y_1, ..., y_j \rangle$ entrambi diversi dalla sequenza vuota che terminano con due elementi diversi.

Se i due prefissi terminano con due elementi diversi ($x_i \neq y_j$) allora la lunghezza di una qualunque tra le più lunghe sottosequenze comuni a X_i e Y_j è maggiore o uguale a l se e solo se la lunghezza di una qualunque tra le più lunghe sottosequenze comuni a X_{i-1} e Y_j risulta essere maggiore o uguale l (si considera la soluzione $c_{i-1,j,l}$ relativa al sottoproblema di dimensione $(i-1, j, l)$ con input il prefisso X_{i-1}, il prefisso Y_j e la lunghezza l: si è tolto da X_i l'elemento x_i) oppure la lunghezza di una qualunque tra le più lunghe sottosequenze comuni a X_i e Y_{j-1} risulta essere maggiore o uguale l (si considera la soluzione $c_{i,j-1,l}$ relativa al sottoproblema di dimensione $(i, j-1, l)$ con input il prefisso X_i, il prefisso Y_{j-1} e la lunghezza l: si è tolto da Y_j l'elemento y_j). Infatti, dato che gli elementi terminanti dei due prefissi sono diversi e non possono essere accodati a nessuna soluzione di un sottoproblema di dimensione minore, solo se esiste una più lunga sottosequenza comune a X_{i-1} e Y_j di lunghezza maggiore o uguale a l oppure una più lunga sottosequenza comune a X_i e Y_{j-1} di lunghezza maggiore o uguale a l è possibile affermare che una più lunga sottosequenza comune a X_i e Y_j ha lunghezza maggiore o uguale a l. Il tutto può essere scritto come:

$$c_{i,j,l} = c_{i-1,j,l} \vee c_{i,j-1,l} \qquad \textit{se } i > 0 \wedge j > 0 \wedge l \, 0 \wedge x_i \neq y_j$$

Riassumendo, il passo ricorsivo può essere scritto come:

$$c_{i,j,l} = \begin{cases} c_{i-1,j-1,l-1} & \textit{se } i > 0 \wedge j > 0 \wedge l > 0 \wedge x_i = y_j \\ c_{i-1,j,l} \vee c_{i,j-1,l} & \textit{se } i > 0 \wedge j > 0 \wedge l > 0 \wedge x_i \neq y_j \end{cases}$$

6.4 Soluzione del problema

Una volta calcolati i valori di tutte le variabili $c_{0,0,0}, \ldots, c_{m,n,L}$ è possibile identificare la soluzione al problema `PB`:

$$c_{m,n,L}$$

6.5 Implementazione

L'algoritmo 6 presenta lo pseudocodice che calcola tutti i valori $c_{0,0,0}, \ldots, c_{m,n,L}$ basandosi sulle equazioni di ricorrenza definite nella Sezione 6.3. Anziché utilizzare la ricorsione pura, tutti i valori sono calcolati con tecnica **bottom-up** in modo tale da risolvere ogni sottoproblema una volta sola. Questo permette di risolvere il problema in $\mathcal{O}(m \cdot n \cdot L)$ occupando $\Theta(m \cdot n \cdot L)$ spazio in memoria (ossia una matrice tridimensionale che contiene i vari $c_{i,j,L}$).

Algorithm 6 Esistenza di una più lunga sottosequenza di lunghezza $\geq L$ comune fra due sequenze X e Y di lunghezza m e n rispettivamente.

```
procedure LCS≥L(X, Y, L)
    for i ← 0 to m do
        for j ← 0 to n do
            c[i, j, 0] ← true
    for i ← 0 to m do
        for l ← 1 to L do
            c[i, 0, l] ← false
    for j ← 0 to n do
        for l ← 1 to L do
            c[0, j, l] ← false
    for i ← 1 to m do
        for j ← 1 to n do
            for l ← 1 to L do
                if x_i = y_j then
                    c[i, j, l] ← c[i − 1, j − 1, l − 1]
                else
                    c[i, j, l] ← c[i − 1, j, l] ∨ c[i, j − 1, l]
    return c[m, n, L]
```

7 Più lunga sottosequenza comune di lunghezza al più L (LCS≤L)

7.1 Definizione del problema

Il problema `PB` è definito come segue:

date due sequenze X e Y, rispettivamente di m e n numeri interi, e un naturale L, stabilire se la lunghezza di una qualunque tra le più lunghe sottosequenze comuni a X e Y è minore o uguale a L

Ad esempio, date $X = \langle 2, 4, 7, 11, 14, 21, 8\rangle$, $Y = \langle 2, 7, 4, 23, 11, 14, 21, 1\rangle$ e $L = 4$, la soluzione al problema deve essere *false* in quanto tutte le più lunghe sottosequenze comuni hanno lunghezza 5 (che è maggiore di 4).

Per risolvere il problema `PB` non è necessario introdurre alcuna versione ridotta.

7.2 Sottoproblemi e variabili associate

Il problema `PB` contiene diversi sottoproblemi ognuno dei quali non ha come input la tripla (X, Y, L) ma una tripla composta dai prefissi delle sequenze X e Y e da un certo naturale $0 \leq l \leq L$.

Il **sottoproblema di dimensione** (i, j, l) è definito come segue:

date due sequenze X e Y, rispettivamente di m ed n numeri interi, e un naturale l, stabilire se la lunghezza di una qualunque tra le più lunghe sottosequenze comuni a X_i e Y_j è minore o uguale a l

Dato che $0 \leq i \leq m$, $0 \leq j \leq n$ e $0 \leq l \leq L$, si ottengono $(m+1) \cdot (n+1) \cdot (L+1)$ sottoproblemi (i e j possono valere 0 in quanto si deve considerare anche il caso in cui un prefisso sia la sequenza vuota). Ad ogni sottoproblema di `PBR` è associata una **variabile**.

Considerato il sottoproblema di dimensione (i, j, l), la variabile ad esso associata è $c_{i,j,l}$ ed è così definita:

> $c_{i,j,l} =$ ***true*** *se e solo se la lunghezza di una qualunque tra le più lunghe sottosequenze comuni a* X_i *e* Y_j *è minore o uguale a* l*,* ***false*** *altrimenti*

Per determinare la soluzione di un qualsiasi sottoproblema di dimensione (i, j, l), oltre all'input del problema, si utilizzeranno le soluzioni dei sottoproblemi di dimensione minore. Per esempio, dato il sottoproblema di dimensione $(1, 1, 1)$, il valore della variabile $c_{1,1,1}$ si otterrà utilizzando le variabili $c_{0,0,0}$, $c_{0,0,1}$, $c_{1,0,0}$, $c_{1,0,1}$, $c_{0,1,0}$ e $c_{0,1,1}$. Si noti, però, che ognuna di queste variabili è da considerare come una ***black-box***: si può utilizzare ma non è possibile conoscerne il contenuto.

Considerato un qualunque sottoproblema di dimensione (i, j, l)*, assumendo di aver già risolto tutti i sottoproblemi di dimensione minore (ossia sapendo di poter utilizzare tutte le variabili* $c_{i',j',l'}$ *relative ai sottoproblemi di dimensione minore senza però sapere effettivamente quanto valgono), come è possibile ottenere la soluzione* $c_{i,j,l}$ *andando a combinare le soluzioni dei sottoproblemi di dimensione minore a* (i, j, l) *ed, eventualmente, accodando gli elementi* x_i *e* y_j *estratti dall'input* (X_i, Y_i) *e utilizzando il valore* l*?*

7.3 Equazioni di ricorrenza

Per rispondere a questa domanda, si devono andare a definire le **equazioni di ricorrenza**.

Le equazioni di ricorrenza si scrivono senza sapere il valore delle soluzioni dei sottoproblemi ma sapendo solamente che tali soluzioni esistono e si possono utilizzare

Un'equazione di ricorrenza è composta da:

- un **caso base** che definisce i casi più semplici che possono essere subito risolti senza ricorrere alle soluzioni dei sottoproblemi più piccoli

- un **passo ricorsivo** che definisce come risolvere i casi più complessi a partire dalle soluzioni dei sottoproblemi più piccoli (che si assume essere stati già risolti)

7.3.1 Caso base

Il caso base si ha in due situazioni differenti:

1. per un qualunque sottoproblema di dimensione (i, j, l) con

$$i = 0 \lor j = 0$$

2. per un qualunque sottoproblema di dimensione (i, j, l) con

$$i > 0 \wedge j > 0 \wedge l = 0 \ \textit{tale che} \ x_i = y_j$$

Per entrambi questi casi, infatti, è facile stabilire il valore di $c_{i,j,l}$.

Caso base 1: (i, j, l) **con** $i = 0 \vee j = 0$ Questo caso base si verifica quando uno dei due prefissi considerati è la sequenza vuota. È facile ottenere il valore della variabile $c_{i,j,l}$: per qualsiasi valore di l, $c_{i,j,l}$ vale *true* in quanto la lunghezza della più lunga sottosequenza comune fra la sequenza vuota e una qualsiasi altra sequenza è pari a 0 (che è minore o uguale di qualsiasi valore può assumere l). Questo caso base può essere scritto nel seguente modo:

$$c_{i,j,l} = true \ se \ i = 0 \vee j = 0$$

Caso base 2: (i, j, l) **con** $i > 0 \wedge j > 0 \wedge l = 0$ ***tale che*** $x_i = y_j$

Si capirà che questo è un caso base una volta affrontato il passo ricorsivo 1 (Sezione 7.3.2)

Questo caso base si verifica quando entrambi i prefissi considerati sono diversi dalla stringa vuota, il loro ultimo elemento è uguale e si richiede che tutte le più lunghe sottosequenze comuni abbiano lunghezza minore o uguale a 0. Questo è sempre falso in quanto $x_i = y_j$ e una fra le più lunghe sottosequenze comuni a X_i e Y_j deve contenere per forza l'elemento $x_i = y_j$ e quindi deve avere lunghezza almeno 1 (che non è minore o uguale a $l = 0$). Questo caso base può essere scritto nel seguente modo:

$$c_{i,j,l} = false \quad se \ i > 0 \wedge j > 0 \wedge l = 0 \wedge x_i = y_j$$

Riassumendo, il caso base può essere scritto nel seguente modo:

$$c_{i,j,l} = \begin{cases} true & se \ i = 0 \vee j = 0 \\ false & se \ i > 0 \wedge j > 0 \wedge l = 0 \wedge x_i = y_j \end{cases}$$

7.3.2 Passo ricorsivo

Il passo ricorsivo si ha nelle seguenti situazioni:

1. per un qualunque sottoproblema di dimensione (i, j, l) con

$$i > 0 \wedge j > 0 \wedge l > 0 \wedge x_i = y_j$$

2. per un qualunque sottoproblema di dimensione (i, j, l) con

$$i > 0 \wedge j > 0 \wedge l \geq 0 \wedge x_i \neq y_j$$

Passo ricorsivo 1: (i, j, l) **con** $i > 0 \wedge j > 0 \wedge l > 0$ ***tale che*** $x_i = y_j$
Questo passo ricorsivo si verifica quando si vanno a considerare due prefissi $X_i = \langle x_0, x_1, ..., x_i \rangle$ e $Y_j = \langle y_0, y_1, ..., y_j \rangle$ entrambi diversi dalla sequenza vuota che terminano con lo stesso elemento.

Se i due prefissi terminano con lo stesso elemento ($x_i = y_j$) allora la lunghezza di una tra le più lunghe sottosequenze comuni a X_i e Y_j è minore o uguale a l se e solo se tutte le più lunghe sottosequenze comuni a X_{i-1} e Y_{j-1} hanno lunghezza minore o uguale a $l-1$ (si considera la soluzione $c_{i-1,j-1,l-1}$ relativa al sottoproblema di dimensione $(i-1, j-1, l-1)$ con input il prefisso X_{i-1}, il prefisso Y_{j-1} e la lunghezza $l-1$). Infatti, solo se tutte le più lunghe sottosequenze comuni a X_{i-1} e Y_{j-1} hanno lunghezza minore o uguale a $l-1$ allora è possibile concatenare ad ognuna di esse l'elemento comune $x_i = y_j$ ed ottenere una qualsiasi sottosequenza comune a X_i e Y_j di lunghezza minore o uguale a l. Tutto questo, però, è corretto solamente se $\underline{l > 0}$ in quanto se l fosse uguale a 0, si dovrebbe andare a considerare la soluzione di un sottoproblema di dimensione $(i-1, j-1, -1)$ che non esiste.

Riassumendo, questo passo ricorsivo può essere scritto come:

$$c_{i,j,l} = c_{i-1,j-1,l-1} \qquad \textit{se } i > 0 \wedge j > 0 \wedge l > 0 \wedge x_i = y_j$$

Nel caso in cui $l = 0$, il sottoproblema è facile da risolvere senza dover ricorrere all'utilizzo delle soluzioni dei sottoproblemi di dimensione minore: quando $l = 0$ il caso qui descritto è un caso base (si veda il caso base 2 nella Sezione 7.3.1).

Passo ricorsivo 2: (i, j, l) **con** $i > 0 \wedge j > 0 \wedge l \geq 0$ ***tale che*** $x_i \neq y_j$
Questo passo ricorsivo si verifica quando si vanno a considerare due prefissi $X_i = \langle x_0, x_1, ..., x_i \rangle$ e $Y_j = \langle y_0, y_1, ..., y_j \rangle$ entrambi diversi dalla sequenza vuota che terminano con due elementi diversi. In questo caso, la lunghezza l non è significativa: il valore della soluzione $c_{i,j,l}$ è lo stesso per qualsiasi valore di $l \geq 0$.

Se i due prefissi terminano con due elementi diversi ($x_i \neq y_j$) allora la lunghezza di una tra le più lunghe sottosequenze comuni a X_i e Y_j è minore o uguale a l se e solo se tutte le più lunghe sottosequenze comuni a X_{i-1} e Y_j hanno lunghezza minore o uguale a l (si considera la soluzione $c_{i-1,j,l}$, ossia la soluzione del sottoproblema di dimensione $(i-1, j, l)$ con input il prefisso X_{i-1}, il prefisso Y_j e la lunghezza l: si è tolto da X_i l'elemento x_i) **<u>E</u>** tutte le più lunghe sottosequenze comuni a X_i e Y_{j-1} hanno lunghezza minore o uguale a l (si considera la soluzione $c_{i,j-1,l}$, ossia la soluzione del sottoproblema di dimensione $(i, j-1, l)$ con input il prefisso X_i, il prefisso Y_{j-1} e la lunghezza l: si è tolto da Y_j l'elemento y_j). Infatti, dato che gli elementi terminanti dei due prefissi sono diversi e non possono essere accodati a nessuna soluzione di un sottoproblema di dimensione minore, solo se tutte le più lunghe sottosequenze comuni a X_{i-1} e Y_j e tutte le più lunghe sottosequenze comuni a X_i e Y_{j-1} hanno lunghezza minore o uguale a l allora è possibile affermare che una qualsiasi fra le più lunghe sottosequenze comuni a X_i e Y_j ha lunghezza minore o uguale a l. Il tutto può essere scritto come:

$$c_{i,j,l} = c_{i-1,j,l} \wedge c_{i,j-1,l} \qquad \textit{se } i > 0 \wedge j > 0 \wedge l \geq 0 \wedge x_i \neq y_j$$

Riassumendo, il passo ricorsivo può essere scritto come:

$$c_{i,j,l} = \begin{cases} c_{i-1,j-1,l-1} & se\ i > 0 \wedge j > 0 \wedge l > 0 \wedge x_i = y_j \\ c_{i-1,j,l} \wedge c_{i,j-1,l} & se\ i > 0 \wedge j > 0 \wedge l \geq 0 \wedge x_i \neq y_j \end{cases}$$

7.4 Soluzione del problema

Una volta calcolati i valori di tutte le variabili $c_{0,0,0}, \ldots, c_{m,n,L}$ è possibile identificare la soluzione al problema `PB`:

$$c_{m,n,L}$$

7.5 Implementazione

L'algoritmo 7 presenta lo pseudocodice che calcola tutti i valori $c_{0,0,0}, \ldots, c_{m,n,L}$ basandosi sulle equazioni di ricorrenza definite nella Sezione 7.3. Anziché utilizzare la ricorsione pura, tutti i valori sono calcolati con tecnica **bottom-up** in modo tale da risolvere ogni sottoproblema una volta sola. Questo permette di risolvere il problema in $\mathcal{O}(m \cdot n \cdot L)$ occupando $\Theta(m \cdot n \cdot L)$ spazio in memoria (ossia una matrice tridimensionale che contiene i vari $c_{i,j,L}$).

Algorithm 7 Esistenza di una più lunga sottosequenza di lunghezza $\leq L$ comune fra due sequenze X e Y di lunghezza m e n rispettivamente.

```
procedure LCS≤L(X, Y, L)
    for i ← 0 to m do
        for l ← 0 to L do
            c[i, 0, l] ← true
    for j ← 0 to n do
        for l ← 0 to L do
            c[0, j, l] ← true
    for i ← 1 to m do
        for j ← 1 to n do
            if x_i = y_j then
                c[i, j, 0] ← false
    for i ← 1 to m do
        for j ← 1 to n do
            for l ← 0 to L do
                if x_i = y_j then
                    if l > 0 then
                        c[i, j, l] ← c[i − 1, j − 1, l − 1]
                else
                    c[i, j, l] ← c[i − 1, j, l] ∧ c[i, j − 1, l]
    return c[m, n, L]
```

8 Più lunga sottosequenza comune con al più R elementi rossi (LCSR)

Premessa Sia C un insieme di colori. Gli elementi della sequenza considerata in questo esercizio sono numeri naturali colorati. La funzione di colorazione è definita nel seguente modo: $\phi : \mathbb{N} \rightarrow C$. Per esempio, data la funzione:

$$\phi(x) = \begin{cases} \text{rosso} & x < 5 \\ \text{blu} & 5 \leq x \leq 10 \\ \text{verde} & x > 10 \end{cases}$$

la sequenza $X = \langle 2, 4, 7, 6, 11, 3, 21, 14, 1 \rangle$ sarà colorata nel seguente modo: $\langle r, r, b, b, g, r, g, g, r \rangle$.

8.1 Definizione del problema

Il problema `PB` è definito come segue:

date due sequenze X e Y, rispettivamente di m e n numeri interi, e un naturale R, stabilire se tutte le più lunghe sottosequenze comuni a X e Y contengono al più R elementi colorati di rosso

Ad esempio, dati $X = \langle 2, 7, 3, 11, 21, 14, 1 \rangle$, $Y = \langle 2, 3, 7, 23, 21, 14, 1, 8 \rangle$ e $R = 4$, la soluzione al problema deve essere *true* in quanto tutte le più lunghe sottosequenze comuni ($\langle 2, 7, 21, 14, 1 \rangle$ e $\langle 2, 3, 21, 14, 1 \rangle$) contengono, rispettivamente, 2 e 3 elementi colorati di rosso.

Per risolvere il problema `PB` non è necessario introdurre alcuna versione ridotta.

8.2 Sottoproblemi e variabili associate

Il problema `PB` contiene diversi sottoproblemi ognuno dei quali non ha come input la tripla (X, Y, R) ma una tripla composta dai prefissi delle sequenze X e Y e da un certo naturale $0 \leq r \leq R$.

Il **sottoproblema di dimensione** (i, j, r) è definito come segue:

date due sequenze X e Y, rispettivamente di m ed n numeri interi, e un naturale r, stabilire se tutte le più lunghe sottosequenze comuni a X_i e Y_j contengono al più r elementi colorati di rosso

Dato che $0 \leq i \leq m$, $0 \leq j \leq n$ e $0 \leq r \leq R$, si ottengono $(m+1) \cdot (n+1) \cdot (R+1)$ sottoproblemi (i e j possono valere 0 in quanto si deve considerare anche il caso in cui un prefisso sia la sequenza vuota). Ad ogni sottoproblema di `PBR` è associata una **variabile**.

Considerato il sottoproblema di dimensione (i, j, r), la variabile ad esso associata è $c_{i,j,r}$ ed è così definita:

$c_{i,j,r} =$ ***true*** *se e solo se tutte le più lunghe sottosequenze comuni a X_i e Y_j contengono al più r elementi colorati di rosso,* ***false*** *altrimenti*

Per determinare la soluzione di un qualsiasi sottoproblema di dimensione (i, j, r), oltre all'input del problema, si utilizzeranno le soluzioni dei sottoproblemi di dimensione minore. Per esempio, dato il sottoproblema di dimensione $(1, 1, 1)$, il valore della variabile $c_{1,1,1}$ si otterrà utilizzando le variabili $c_{0,0,0}$, $c_{0,0,1}$, $c_{1,0,0}$, $c_{1,0,1}$, $c_{0,1,0}$ e $c_{0,1,1}$. Si noti, però, che ognuna di queste variabili è da considerare come una ***black-box***: si può utilizzare ma non è possibile conoscerne il contenuto.

Considerato un qualunque sottoproblema di dimensione (i, j, r), assumendo di aver già risolto tutti i sottoproblemi di dimensione minore (ossia sapendo di poter utilizzare tutte le variabili $c_{i',j',r'}$ relative ai sottoproblemi di dimensione minore senza però sapere effettivamente quanto valgono), come è possibile ottenere la soluzione $c_{i,j,r}$ andando a combinare le soluzioni dei sottoproblemi di dimensione minore a (i, j, r) ed, eventualmente, accodando gli elementi x_i e y_j estratti dall'input (X_i, Y_i) e utilizzando il valore r?

8.3 Equazioni di ricorrenza

Per rispondere a questa domanda, si devono andare a definire le **equazioni di ricorrenza**.

Le equazioni di ricorrenza si scrivono senza sapere il valore delle soluzioni dei sottoproblemi ma sapendo solamente che tali soluzioni esistono e si possono utilizzare

Un'equazione di ricorrenza è composta da:

- un **caso base** che definisce i casi più semplici che possono essere subito risolti senza ricorrere alle soluzioni dei sottoproblemi più piccoli
- un **passo ricorsivo** che definisce come risolvere i casi più complessi a partire dalle soluzioni dei sottoproblemi più piccoli (che si assume essere stati già risolti)

8.3.1 Caso base

Il caso base si ha in due situazioni differenti:

1. per un qualunque sottoproblema di dimensione (i, j, r) con

$$i = 0 \lor j = 0$$

2. per un qualunque sottoproblema di dimensione (i, j, r) con

$$i > 0 \land j > 0 \land r = 0 \textit{ tale che } x_i = y_j \land \phi(x_i) = \phi(y_j) = rosso$$

Per entrambi questi casi, infatti, è facile stabilire il valore di $c_{i,j,r}$.

Caso base 1: (i, j, r) **con** $i = 0 \vee j = 0$ Questo caso base si verifica quando uno dei due prefissi considerati è la sequenza vuota. È facile ottenere il valore della variabile $c_{i,j,r}$: dato che si richiede che tutte le più lunghe sottosequenze comuni contengano al più r elementi colorati di rosso, l'unica più lunga sottosequenza comune fra la sequenza vuota e una qualsiasi sequenza è la sequenza vuota stessa che non contiene alcun elemento e quindi, dato che contiene 0 elementi colorati di rossi, certamente contiene anche al più r elementi colorati di rosso (qualsiasi sia r). Questo caso base può essere scritto nel seguente modo:

$$c_{i,j,r} = true \quad se\ i = 0 \vee j = 0$$

Caso base 2: (i, j, r) **con** $i > 0 \wedge j > 0 \wedge r = 0$ ***tale che*** $x_i = y_j \wedge \phi(x_i) = \phi(y_j) = rosso$

Si capirà che questo è un caso base una volta affrontato il passo ricorsivo 1 (Sezione 8.3.2)

Questo caso base si verifica quando entrambi i prefissi considerati sono diversi dalla stringa vuota, il loro ultimo elemento è uguale e colorato di rosso ma si richiede che tutte le più lunghe sottosequenze comuni contengano al più 0 elementi colorati di rosso. Questo è impossibile in quanto $x_i = y_j$ e almeno una più lunga sottosequenza comune a X_i e Y_j deve contenere per forza questo elemento che è colorato di rosso. Pertanto, non tutte le più lunghe sottosequenze comuni a X_i e Y_j contengono almeno 1 elemento colorato di rosso: il valore della variabile $c_{i,j,r}$ deve essere false. Questo caso base può essere scritto nel seguente modo:

$$c_{i,j,r} = false \quad se\ i > 0 \wedge j > 0 \wedge r = 0 \wedge x_i = y_j \wedge \phi(x_i) = \phi(y_j) = rosso$$

Riassumendo, il caso base può essere scritto nel seguente modo:

$$c_{i,j,r} = \begin{cases} true & se\ i = 0 \vee j = 0 \\ false & se\ i > 0 \wedge j > 0 \wedge r = 0 \wedge x_i = y_j \wedge \phi(x_i) = \phi(y_j) = rosso \end{cases}$$

8.3.2 Passo ricorsivo

Il passo ricorsivo si ha nelle seguenti situazioni:

1. per un qualunque sottoproblema di dimensione (i, j, r) con

$$i > 0 \wedge j > 0 \wedge r > 0\ tale\ che\ x_i = y_j \wedge \phi(x_i) = \phi(y_j) = rosso$$

2. per un qualunque sottoproblema di dimensione (i, j, r) con

$$i > 0 \wedge j > 0 \wedge r \geq 0\ tale\ che\ x_i = y_j \wedge \phi(x_i) = \phi(y_j) \neq rosso$$

3. per un qualunque sottoproblema di dimensione (i, j, r) con

$$i > 0 \wedge j > 0 \wedge r \geq 0\ tale\ che\ x_i \neq y_j$$

Passo ricorsivo 1: (i, j, l) **con** $i > 0 \wedge j > 0$ ***tale che*** $x_i = y_j \wedge \phi(x_i) = \phi(y_j) = rosso$ Questo passo ricorsivo si verifica quando si vanno a considerare due prefissi $X_i = \langle x_0, x_1, ..., x_i\rangle$ e $Y_j = \langle y_0, y_1, ..., y_j\rangle$ entrambi diversi dalla sequenza vuota che terminano con lo stesso elemento colorato di rosso.

Se i due prefissi terminano con lo stesso elemento ($x_i = y_j$) e questo è colorato di rosso allora tutte le più lunghe sottosequenze comuni a X_i e Y_j contengono al più r elementi colorati di rosso se e solo se tutte le più lunghe sottosequenze comuni a X_{i-1} e Y_{j-1} contengono al più $r - 1$ elementi colorati di rosso (si considera la soluzione $c_{i-1,j-1,r-1}$ relativa al sottoproblema di dimensione $(i-1, j-1, r-1)$ con input il prefisso X_{i-1}, il prefisso Y_{j-1} e $r-1$). Infatti, solo se tutte le sottosequenze più lunghe comuni a X_{i-1} e Y_{j-1} contengono al più $r - 1$ elementi colorati di rosso allora è possibile concatenare ad esse l'elemento comune $x_i = y_j$ ed ottenere una qualsiasi più lunga sottosequenza comune a X_i e Y_j con al più r elementi colorati di rosso. Tutto questo, però, è corretto solamente se $\underline{r > 0}$ in quanto se r fosse uguale a 0, si dovrebbe andare a considerare la soluzione di un sottoproblema di dimensione $(i - 1, j - 1, -1)$ che non esiste.

Riassumendo, questo passo ricorsivo può essere scritto come:

$$c_{i,j,r} = c_{i-1,j-1,r-1} \textit{ se } i > 0 \wedge j > 0 \wedge r > 0 \wedge x_i = y_j \wedge \phi(x_i) = \phi(y_j) = rosso$$

Nel caso in cui $r = 0$, il sottoproblema è facile da risolvere senza dover ricorrere all'utilizzo delle soluzioni dei sottoproblemi di dimensione minore: quando $r = 0$ il caso qui descritto è un caso base (si veda il caso base 2 nella Sezione 8.3.1).

Passo ricorsivo 2: (i, j, l) **con** $i > 0 \wedge j > 0 \wedge r \geq 0$ ***tale che*** $x_i = y_j \wedge \phi(x_i) = \phi(y_j) \neq rosso$ Questo passo ricorsivo si verifica quando si vanno a considerare due prefissi $X_i = \langle x_0, x_1, ..., x_i\rangle$ e $Y_j = \langle y_0, y_1, ..., y_j\rangle$ entrambi diversi dalla sequenza vuota che terminano con lo stesso elemento colorato non di rosso. In questo caso, il valore r non è significativo (è necessario che $r \geq 0$): il valore della soluzione $c_{i,j,r}$ è lo stesso per qualsiasi valore di r.

Se i due prefissi terminano con lo stesso elemento che non è però colorato di rosso allora tutte le più lunghe sottosequenze comuni a X_i e Y_j contengono al più r elementi colorati di rosso se e solo se tutte le più lunghe sottosequenze comuni a X_{i-1} e Y_{j-1} contengono al più r elementi colorati di rosso (si considera la soluzione $c_{i-1,j-1,r}$, ossia la soluzione del sottoproblema di dimensione $(i - 1, j - 1, r)$ con input il prefisso X_{i-1}, il prefisso Y_{j-1} e il valore r). Infatti, solo se tutte le più lunghe sottosequenze comuni a X_{i-1} e Y_{j-1} contengono al più r elementi colorati di rosso allora è possibile concatenare a qualunque sottosequenza di esse l'elemento comune $x_i = y_j$ e poter affermare ancora che tutte le più lunghe sottosequenze comuni a X_i e Y_j contengono al più r elementi colorati di rosso (si ricorda che l'elemento da concatenare non è colorato di rosso). Il tutto può essere scritto come:

$$c_{i,j,r} = c_{i-1,j-1,r} \textit{ se } i > 0 \wedge j > 0 \wedge r \geq 0 \wedge x_i = y_j \wedge \phi(x_i) = \phi(y_j) \neq rosso$$

Passo ricorsivo 3: (i, j, l) **con** $i > 0 \wedge j > 0 \wedge r \geq 0$ ***tale che*** $x_i \neq y_j$ Questo passo ricorsivo si verifica quando si vanno a considerare due prefissi

$X_i = \langle x_0, x_1, ..., x_i \rangle$ e $Y_j = \langle y_0, y_1, ..., y_j \rangle$ entrambi diversi dalla sequenza vuota che terminano con due elementi differenti. In questo caso, il colore degli elementi e il valore r non sono significativi (è necessario che $r \geq 0$): il valore della soluzione $c_{i,j,r}$ è lo stesso per qualsiasi colore e per qualsiasi valore di r.

Se i due prefissi terminano con due elementi differenti, tutte le più lunghe sottosequenze comuni a X_i e Y_j contengono al più r elementi colorati di rosso se e solo se tutte le più lunghe sottosequenze comuni a X_{i-1} e Y_j contengono al più r elementi colorati di rosso (si considera la soluzione $c_{i-1,j,r}$, ossia la soluzione del sottoproblema di dimensione $(i-1, j, r)$ con input il prefisso X_{i-1}, il prefisso Y_j e il valore r: si è tolto da X_i l'elemento x_i) e tutte le più lunghe sottosequenze comuni a X_i e Y_{j-1} contengono al più r elementi colorati di rosso (si considera la soluzione $c_{i,j-1,r}$, ossia la soluzione del sottoproblema di dimensione $(i, j-1, r)$ con input il prefisso X_i, il prefisso Y_{j-1} e il valore r: si è tolto da Y_j l'elemento y_j). Infatti, solo se tutte le più lunghe sottosequenze comuni a X_{i-1} e Y_j contengono al più r elementi colorati di rosso e tutte le più lunghe sottosequenze comuni a X_i e Y_{j-1} contengono al più r elementi colorati di rosso allora tutte le più lunghe sottosequenze comuni a X_i e Y_j contengono al più r elementi colorati di rosso. Il tutto può essere scritto come:

$$c_{i,j,r} = c_{i-1,j,r} \land c_{i,j-1,r} \qquad se\ i > 0 \land j > 0 \land r \geq 0 \land x_i \neq y_j$$

Riassumendo, il passo ricorsivo può essere scritto come:

$$c_{i,j,r} = \begin{cases} c_{i-1,j-1,r-1} & se\ i > 0 \land j > 0 \land r > 0 \land x_i = y_j \land \phi(x_i) = \phi(y_j) = rosso \\ c_{i-1,j-1,r} & se\ i > 0 \land j > 0 \land r \geq 0 \land x_i = y_j \land \phi(x_i) = \phi(y_j) \neq rosso \\ c_{i-1,j,r} \land c_{i,j-1,r} & se\ i > 0 \land j > 0 \land r \geq 0 \land x_i \neq y_j \end{cases}$$

8.4 Soluzione del problema

Una volta calcolati i valori di tutte le variabili $c_{0,0,0}, \ldots, c_{m,n,R}$ è possibile identificare la soluzione al problema `PB`:

$$c_{m,n,R}$$

8.5 Implementazione

L'algoritmo 8 presenta lo pseudocodice che calcola tutti i valori $c_{0,0,0}, \ldots, c_{m,n,R}$ basandosi sulle equazioni di ricorrenza definite nella Sezione 8.3. Anziché utilizzare la ricorsione pura, tutti i valori sono calcolati con tecnica **bottom-up** in modo tale da risolvere ogni sottoproblema una volta sola. Questo permette di risolvere il problema in $\mathcal{O}(m \cdot n \cdot R)$ occupando $\Theta(m \cdot n \cdot R)$ spazio in memoria (ossia una matrice tridimensionale che contiene i vari $c_{i,j,R}$).

Algorithm 8 Esistenza di una più lunga sottosequenza comune fra due sequenze X e Y di lunghezza m e n rispettivamente con al più R elementi colorati di rosso.

```
procedure LCSR(X, Y, R)
    for i ← 0 to m do
        for r ← 0 to R do
            c[i, 0, r] ← true
    for j ← 0 to n do
        for r ← 0 to R do
            c[0, j, r] ← true
    for i ← 1 to m do
        for j ← 1 to n do
            if x_i = y_j and φ(x_i) = rosso then
                c[i, j, 0] ← false
    for i ← 1 to m do
        for j ← 1 to n do
            for r ← 0 to R do
                if x_i = y_j then
                    if φ(x_i) = rosso then
                        if r > 0 then
                            c[i, j, r] ← c[i − 1, j − 1, r − 1]
                    else
                        c[i, j, r] ← c[i − 1, j − 1, r]
                else
                    c[i, j, r] ← c[i − 1, j, r] ∧ c[i, j − 1, r]
    return c[m, n, R]
```