Laboratorio di Algoritmi e Strutture Dati

Docenti: M. Torelli, S. Aguzzoli

Appello di settembre 2009

Progetto "Colori"

Consegna entro mercoledì 23 settembre 2009

Il problema

Gli abitanti di una remota isola dell'Oceno Pacifico soffrono di una strana forma di disturbo della percezione dei colori. Per studiare il fenomeno, un'equipe di medici sottopone loro dei test allo scopo di controllare quali colori vengano confusi.

Un test è costituito da una tavola di dimensione imprecisata divisa in celle quadrate di dimensione unitaria. Ogni cella può essere colorata con un singolo colore.

Formalmente, la tavola è identificata con il piano

$$\mathbb{Z} \times \mathbb{Z} \;\; = \;\; \{\, (x, y) \;\mid\; x \in \mathbb{Z},\, y \in \mathbb{Z} \,\}$$

delle coppie di interi. Ogni coppia di interi $(x, y) \in \mathbb{Z} \times \mathbb{Z}$ identifica una *cella*. La cella (x, y) è *adiacente* alle celle:

$$(x - 1, y),\; (x + 1, y),\; (x, y - 1),\; (x, y + 1)\,.$$

Un *colore* è una stringa di lunghezza arbitraria sull'alfabeto $a, b, \ldots, z$ delle lettere minuscole. Se la cella (x, y) è colorata, $C(x, y)$ ne denota il colore.

Sia $\mathcal{R}$ un insieme di coppie di colori. Diciamo che due colori α e β sono *equivalenti* rispetto a $\mathcal{R}$ (in simboli, $\alpha \equiv_{\mathcal{R}} \beta$) se e solo se esiste una sequenza di colori γ_1,. . . γ_n ($n \geq 1$) tali che:

- $\alpha = \gamma_1$ e $\beta = \gamma_n$;
- per ogni $i \in \{1, \ldots, n - 1\}$, $(\gamma_i, \gamma_{i+1}) \in \mathcal{R}$ oppure $(\gamma_{i+1}, \gamma_i) \in \mathcal{R}$.

In altri termini, $\equiv_{\mathcal{R}}$ è la più piccola relazione di equivalenza contenente la relazione $\mathcal{R}$. La relazione $\equiv_{\mathcal{R}}$ contiene tutte e sole le coppie di colori che un paziente σ non è in grado di distinguere.

Esempio 1

Supponiamo che l'insieme $\mathcal{R}$ contenga le coppie:

$$(\mathit{giallo}, \mathit{verde}),\;\; (\mathit{rosso}, \mathit{blu}),\;\; (\mathit{blu}, \mathit{nero}),\;\; (\mathit{verde}, \mathit{azzurro}),\;\; (\mathit{blu}, \mathit{viola})$$

Valgono, ad esempio, le seguenti relazioni:

$\mathit{verde} \equiv_{\mathcal{R}} \mathit{giallo}$ $\quad$ $\mathit{azzurro} \equiv_{\mathcal{R}} \mathit{giallo}$ $\quad$ $\mathit{nero} \equiv_{\mathcal{R}} \mathit{rosso}$ $\quad$ $\mathit{viola} \equiv_{\mathcal{R}} \mathit{rosso}$ $\quad$ $\mathit{rosso} \equiv_{\mathcal{R}} \mathit{nero}$

Rispetto a $\mathcal{R}$, *rosso* e *verde* non sono equivalenti, e neppure *azzurro* e *rosso*.

Dato un insieme finito $\mathcal{R}$ di coppie di colori, due celle colorate (x', y') e (x'', y'') sono *connesse* rispetto a $\mathcal{R}$ se esiste una sequenza di celle colorate $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ tale che:

1. $(x_1, y_1) = (x', y')$ e $(x_n, y_n) = (x'', y'')$,

2. (x_i, y_i) è adiacente a (x_{i+1}, y_{i+1}) per ogni $i \in \{1, \ldots, n-1\}$.

3. Per ogni coppia di celle (x_i, y_i) e (x_j, y_j) della sequenza, $C(x_i, y_i) \equiv_{\mathcal{R}} C(x_j, y_j)$.

Una *figura* del piano definita da $\mathcal{R}$ è un insieme massimale $\mathcal{M} \subset \mathbb{Z} \times \mathbb{Z}$ di celle colorate connesse rispetto ad $\mathcal{R}$: vale a dire che, se $\mathcal{N}$ è un altro insieme di celle colorate connesse rispetto ad $\mathcal{R}$ e $\mathcal{M} \cap \mathcal{N} \neq \emptyset$ allora $\mathcal{N} \subseteq \mathcal{M}$. Una figura definita da $\mathcal{R}$ è quindi un insieme massimale di celle connesse che appaiono dello stesso colore a un osservatore che non sa distinguere i colori in $\mathcal{R}$.

Esempio 2
Supponiamo che le celle siano colorate come in Figura 1, dove A rappresente il colore *azzurro*, G il *giallo*, R il *rosso* e V il *verde* (le celle vuote non sono colorate). In tale configurazione vale:

$$C(1,4) = \textit{giallo} \qquad C(2,2) = \textit{verde} \qquad C(2,4) = \textit{giallo} \qquad C(3,2) = \textit{giallo} \qquad \ldots$$

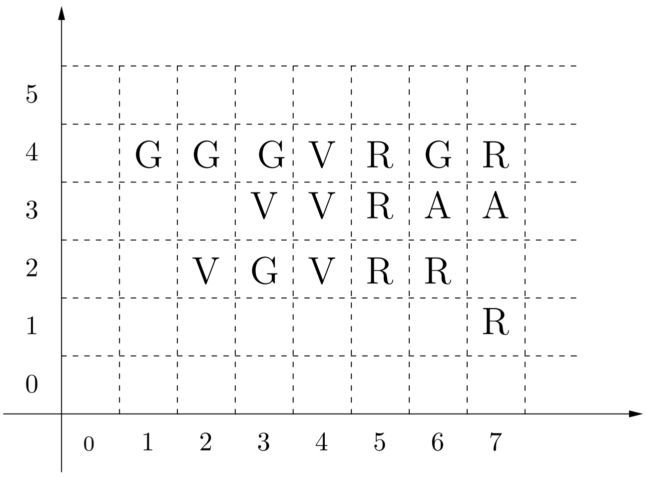

Figura 1: Esempio di colorazione di celle

Sia $\mathcal{R}_0$ l'insieme vuoto. Le figure definite da $\mathcal{R}_0$ sono nove:

$$\begin{aligned}
&F_1 = \{(1,4), (2,4), (3,4)\}, \qquad F_2 = \{(2,2)\}, \qquad F_3 = \{(3,2)\}, \\
&F_4 = \{(3,3), (4,2), (4,3), (4,4)\}, \qquad F_5 = \{(5,2), (5,3), (5,4), (6,2)\}, \\
&F_6 = \{(6,3), (7,3)\}, \qquad F_7 = \{(6,4)\}, \qquad F_8 = \{(7,1)\}, \qquad F_9 = \{(7,4)\}.
\end{aligned}$$

Sia $\mathcal{R}_1 = \{(\textit{giallo}, \textit{verde})\}$. Le figure definite da $\mathcal{R}_1$ sono sei:

$$F_{10} = F_1 \cup F_2 \cup F_3 \cup F_4, \qquad F_5, \qquad F_6, \qquad F_7, \qquad F_8, \qquad F_9.$$

Consideriamo ora l'insieme $\mathcal{R}_2 = \{(\textit{giallo}, \textit{verde}), (\textit{rosso}, \textit{azzurro})\}$. Le figure definite da $\mathcal{R}_2$ sono quattro:

$$F_{10}, \qquad F_{11} = F_5 \cup F_6 \cup F_9, \qquad F_7, \qquad F_8.$$

Infine, se $\mathcal{R}_3 = \{(\textit{giallo}, \textit{verde}), (\textit{rosso}, \textit{azzurro}), (\textit{giallo}, \textit{rosso})\}$, tutti i colori in figura diventano equivalenti. Le figure definite da $\mathcal{R}_3$ sono due: $\{(7,1)\}$ e la figura formata dalle rimanenti celle colorate.

Ogni *paziente* è identificato attraverso il suo nome, che è una stringa σ di lunghezza arbitraria sull'alfabeto $a, b, \ldots, z$ delle lettere minuscole. A ogni paziente σ è associato l'insieme $\mathcal{R}_\sigma$ dei colori che egli non è in grado di distinguere (che può essere anche vuoto). Una figura F è riconosciuta da σ se e solo se F è una figura definita da $\mathcal{R}_\sigma$.

Si richiede di implementare una struttura dati efficiente che permetta di eseguire le operazioni seguenti:

- **input** $(r, c, x, y, nomefile)$

  Legge dal file $nomefile$ una tabella di colori di r righe e c colonne e assegna i colori corrispondenti alle celle nell'insieme $\{(x+h, y+k) \mid 0 \leq h < c, 0 \leq k < r\}$ secondo le regole specificate nell'apposita sezione.

- **modifica** (x, y, α)

  Pone $C(x, y) = \alpha$.

- **confonde** (σ, α, β)

  Aggiunge la coppia di colori (α, β) all'insieme $\mathcal{R}_\sigma$ (quindi, dichiara che il paziente σ non distingue i colori α e β).

- **pazienti** (α, β)

  Stampa, in ordine alfabetico, la lista dei pazienti che confondono i colori α e β.

- **test** (σ)

  Stampa il numero delle figure riconosciute da σ.

- **figura** (x, y, σ)

  Se (x, y) è una cella colorata, stampa la figura riconosciuta da σ contente (x, y) secondo il formato specificato nell'apposita sezione. Altrimenti, stampa su una nuova linea il messaggio

  ```
  (x,y) non colorata
  ```

All'inizio del programma il piano è vuoto e $\mathcal{R}_\sigma = \emptyset$ per ogni possibile σ. Si noti che le operazioni richieste sono liberamente implementabili; in particolare, non vanno necessariamente intese come prototipi di funzioni.

Specifiche di implementazione

Il programma deve leggere dallo standard input (`stdin`) una sequenza di righe (separate da `\n`), ciascuna delle quali corrisponde a una riga della prima colonna della Tabella 1, dove $nomefile$ è il nome di un file, α, β e σ sono stringhe finite sull'alfabeto $a, b, \ldots, z$ di lunghezza *arbitraria*, r, c sono interi positivi e x, y sono interi. I vari elementi sulla riga sono separati da uno o più spazi. Quando una riga è letta, viene eseguita l'operazione associata; le operazioni di stampa sono effettuate sullo standard output (`stdout`), e ogni operazione deve iniziare su una nuova riga.

RIGA DI INPUT	OPERAZIONE
`i` $r\ c\ x\ y\ nomefile$	**input** $(r, c, x, y, nomefile)$
`m` $x\ y\ \alpha$	**modifica**(x, y, α)
`c` $\sigma\ \alpha\ \beta$	**confonde** (σ, α, β)
`p` $\alpha\ \beta$	**pazienti** (α, β)
`t` σ	**test** (σ)
`F` $x\ y\ \sigma$	**figura** (x, y, σ)
`f`	Termina l'esecuzione del programma

Tabella 1: Specifiche del programma

Note

1. Non devono essere presenti vincoli sul numero di colori, pazienti e celle colorate, e sulla lunghezza dei nomi di colori e pazienti (se non quelli determinati dal tipo di dato intero). Non si richiede – anzi si sconsiglia – l'uso di grafica, se non per test personali: in modo particolare, non si usi `conio.h` e neppure `clrscr()`.

2. Per semplicità si suppone che l'input sia sempre conforme alle specifiche di Tabella 1, per cui non è necessario controllare la correttezza dell'input. Per leggere l'input si usino le funzioni standard ANSI C `getchar()` e/o `scanf()`.

3. **Specifiche per la lettura di una tabella di colori da un file**

   Si consideri il comando:

   `i` $r\ c\ x\ y\ nomefile$.

   Allora il file di nome $nomefile$ contiene $r \times c$ stringhe finite sull'alfabeto $\{a, b, \ldots, z\} \cup \{0\}$.

   I colori delle celle nel piano vengono modificati in questo modo: Per ogni $0 \leq h < c$ e ogni $0 \leq k < r$, se la stringa $\alpha_{h,k}$ in posizione $k \cdot c + (h+1)$ nel file è diversa da `0` allora si pone $C(x+h, y+k) = \alpha_{h,k}$; se invece la stringa $\alpha_{h,k}$ in posizione $k \cdot c + (h + 1)$ nel file è uguale a `0` non si modifica il colore della cella $(x + h, y + k)$. Si assume che la prima stringa nel file sia in posizione 1 e l'ultima sia in posizione $r \cdot c$.

   Ad esempio, se il comando è `i 2 3 4 5 tabella.txt`, e il contenuto del file `tabella.txt` è il seguente:

   ```
   blu rosso 0
   bianco 0 nero
   ```

   allora i colori delle celle vengono modificati in questo modo:

   $C(4, 5) = blu, C(5, 5) = rosso, C(4, 6) = bianco, C(6, 6) = nero.$

   Nessun'altra cella viene modificata. In particolare le celle $(6, 5)$ e $(5, 6)$ non vengono toccate, in quanto le posizioni a loro corrispondenti in `tabella.txt` contengono la stringa `0`.

4. **Formato per la stampa di una figura**

   Si consideri il comando:

   F x y σ,

   e sia $\{(x_1, y_1), (x_2, y_2), \ldots, (x_u, y_u)\}$ la figura del piano definita da $\mathcal{R}_\sigma$ e contenente (x, y). Allora l'output del comando deve essere visualizzato come segue:

   $(\sigma, x, y$
   x_1, y_1
   x_2, y_2
   $\vdots$
   x_u, y_u
   $)$

   L'ordine in cui vengono elencate le celle della figura è arbitrario.

Esempio
Si supponga che le righe di input siano:

```
i 3 2 -2 1 f1.txt
i 4 4 1 1 f2.txt
i 3 2 2 -3 f3.txt
c barbara giallo rosso
c carla giallo blu
c carla giallo rosso
c daniele arancione blu
c daniele rosso verde
c daniele  rosso giallo
c enrico  blu rosso
c enrico  verde rosso
c francesco arancione blu
c francesco blu giallo
c francesco rosso giallo
c francesco rosso verde
p rosso giallo
c enrico  giallo verde
t anna
F 2 -3 anna
t barbara
t carla
t daniele
t enrico
F 3 3 enrico
t francesco
m -2 2 blu
m 0 1 rosso
m 1 1 rosso
m 2 0 giallo
m 2 -1 giallo
t anna
```

```
t barbara
t carla
F -2 3 carla
t daniele
t enrico
F -2 3 enrico
t francesco
f
```

dove il file `f1.txt` contiene

```
rosso rosso
giallo 0
blu 0
```

`f2.txt` contiene

```
0 giallo 0 0
arancione 0 giallo giallo
arancione blu rosso giallo
0 verde verde 0
```

e `f3.txt` contiene

```
blu blu
blu 0
0 verde
```

L'output prodotto dal programma deve essere il seguente

```
barbara
carla
daniele
francesco
11
(anna,2,-3
2,-2
2,-3
3,-3
)
9
7
7
6
(enrico,3,3
2,3
2,4
3,2
3,3
3,4
4,2
```

```
4,3
)
5
10
8
5
(carla,-2,3
-2,3
-2,2
-2,1
-1,1
0,1
1,1
2,1
2,0
2,-1
2,-2
2,-3
3,-3
)
5
3
(enrico,-2,3
-2,3
-2,2
-2,1
-1,1
0,1
1,1
2,1
2,0
2,-1
3,-1
2,-2
2,-3
3,-3
)
1
```

Presentazione del progetto

Il progetto deve essere inviato per posta elettronica all'indirizzo `aguzzoli@dsi.unimi.it` entro il 23 settembre 2009 (incluso). La discussione del progetto e l'esame orale si svolgeranno in data e luogo da specificarsi (consultare al riguardo il sito: `http://homes.dsi.unimi.it/~aguzzoli/algo.htm`).

Occorre presentare:

1. il codice sorgente (rigorosamente ANSI C, compilabile con **gcc**);
2. una sintetica relazione (formato pdf o rtf) che illustra le strutture dati utilizzate e analizza il costo delle diverse operazioni richieste dalla specifica.

I due o più file (file sorgenti C + relazione) devono essere contenuti in un unico file `.zip` il cui nome dovrà essere `cognome.zip`. La relazione e il codice devono riportare il vostro nome, cognome e matricola.

Una copia cartacea della relazione e del codice deve inoltre essere consegnata al dr. Aguzzoli entro il 23 settembre 2009 (lasciandola eventualmente nella sua casella postale presso il dipartimento in via Comelico).

Si ricorda infine di presentarsi alla prova orale con una copia stampata della relazione e del codice.

Il calendario degli esami orali sarà disponibile sulla pagina del corso
`http://homes.dsi.unimi.it/~aguzzoli/algo.htm`)
qualche giorno dopo il termine di consegna del progetto.

Per ogni ulteriore chiarimento:
E-mail: `aguzzoli@dsi.unimi.it`
Ricevimento: il mercoledì, ore 15-16, stanza S204.

Avvisi

La versione aggiornata del progetto è pubblicata in .pdf sul sito:
`http://homes.dsi.unimi.it/~aguzzoli/algo.htm`.
Si consiglia di consultare periodicamente questo sito per eventuali correzioni e/o precisazioni relative al testo del progetto.

Si richiede allo studente di effettuare un adeguato collaudo del proprio progetto su numerosi esempi diversi per verificarne la correttezza e valutarne le prestazioni.

La realizzazione del progetto è una prova d'esame da svolgersi **individualmente**. I progetti giudicati frutto di **collaborazioni** saranno **estromessi** d'ufficio dalla valutazione.