

Esercizio

Si consideri il seguente schema relazionale

AssociazioneCulturale(idSocio,nome,cognome,dataNascita,dataIscrizione,quota,hobby)

Si considerino inoltre le seguenti interrogazioni che vengono effettuate di sovente su tale schema:

```
SELECT cognome, nome, dataNascita
FROM AssociazioneCulturale
ORDERBY Cognome
```

```
SELECT cognome, nome, quota
FROM AssociazioneCulturale
WHERE quota BETWEEN 100 AND 300 AND hobby="Thrillers"
```

```
SELECT cognome, nome, hobby
FROM AssociazioneCulturale
WHERE cognome='Rossi' AND quota BETWEEN 150 AND 500
```

1. Presentare una struttura di memorizzazione primaria della relazione e strutture ausiliarie di accesso che permettano di ottimizzare l'esecuzione delle precedenti interrogazioni.

Controllo ogni interrogazione e vedo cosa e' meglio per ognuna di esse. Dopodiché faccio una scelta che mi permetta di ottimizzare tutte e tre.

INTERROGAZIONE 1

L'ordinamento su Cognome suggerisce di ordinare le tuple della relazione rispetto a tale attributo. Non e' necessaria nessuna struttura ausiliaria di accesso dal momento che non seleziono nessuna tupla.

INTERROGAZIONE 2

Si potrebbero ordinare le tuple della relazione rispetto all'attributo quota. Dopodiché si possono creare un indice B+-tree rispetto a quota e un indice Hash rispetto a hobby. (potrebbe essere sufficiente un solo indice, ma questo dipende dalla selettività dei predicati e nel testo non viene detto nulla).

NOTA: l'organizzazione primaria discussa per questa interrogazione e' incompatibile con l'organizzazione primaria proposta per l'interrogazione 1. (da tenere in considerazione per la scelta finale).

INTERROGAZIONE 3

Si potrebbero ordinare le tuple della relazione rispetto all'attributo quota. Dopodiché si possono creare un indice B+-tree rispetto a quota e un indice Hash rispetto a cognome. (potrebbe essere sufficiente un solo indice, ma questo dipende dalla selettività dei predicati e nel testo non viene detto nulla).

NOTA: : l'organizzazione primaria discussa per questa interrogazione e' compatibile con quella proposta per l'interrogazione 2. Entrambe sono incompatibile con l'organizzazione primaria proposta per l'interrogazione 1. (da tenere in considerazione per la scelta finale).

SCELTA FINALE:

Dall'analisi effettuata per le tre interrogazioni risulta che da una parte si vorrebbero organizzare le tuple rispetto all'attributo cognome, mentre dall'altra si vorrebbe organizzare i dati rispetto all'attributo quota. Dal momento che la prima interrogazione deve accedere tutti i dati ed effettuare un ordinamento di una relazione con molte tuple può essere un'operazione molto costosa, si preferisce scegliere come organizzazione primaria di ordinare le tuple rispetto a cognome.

Dopodiché si può creare un indice B+-tree su quota al fine di ottimizzare la seconda e la terza interrogazione. Infine, si possono creare un indice Hash su cognome e un indice hash su hobby. Le strutture ad indici devono essere dense e non clusterizzate.

NOTA: se ci fossero state solo l'interrogazione 1 e 3, dal momento che sebbene Rossi sia un cognome frequente e' difficile che tutti i membri della associazione culturale si chiamano Rossi, era meglio creare un indice hash su cognome e non creare l'indice B+-tree su quota.

2. Fare vedere un esempio di interrogazione la cui esecuzione viene penalizzata sulla base della organizzazione fisica presentata al punto 1

```
SELECT nome
FROM AssociazioneCulturale
WHERE dataNascita = '09/12/1972'
```

Questa interrogazione e' costosa perche' non sono state definite strutture di accesso al campo dataNascita.