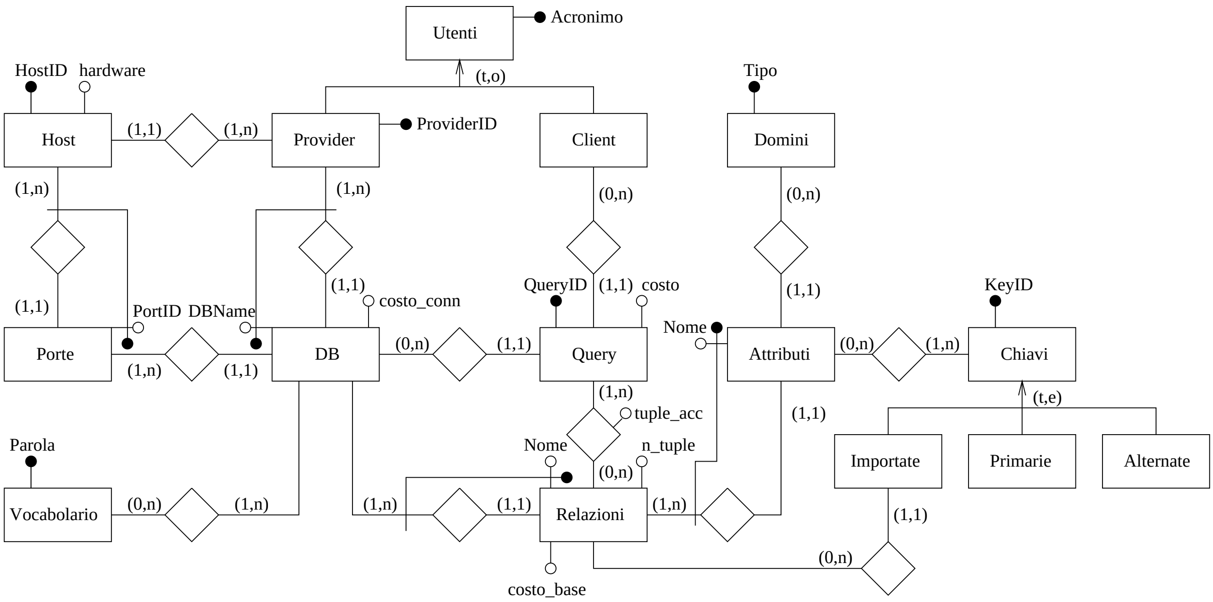
Corso di Laurea in Informatica

Basi di Dati e Sistemi Informativi – 19 Luglio 1997

Risoluzione

Esercizio 1



Nello schema in figura si distinguono tre parti, una che modella aspetti legati principalmente ai `Provider`, la seconda che descrive le strutture dei `DB` e le loro caratteristiche, e la terza che riguarda i `Client`.

Gli aspetti più salienti dello schema sono così riassumibili:

- L'associazione tra `Provider` e `Host` può essere eliminata, in quanto ridondante. Infatti gli `Host` di un `Provider` sono ricavabili conoscendo i suoi `DB` e le `Porte` su cui sono accessibili.

- Un'osservazione analoga vale per l'associazione tra `Query` e `DB`, in quanto tutte le `Relazioni` interessate da una `Query` fanno parte di uno stesso `DB`.

- L'attributo `costo` dell'entità `Query`, che rappresenta il costo complessivo della query, è un attributo derivato, in quanto i suoi valori sono calcolabili a partire da quelli di altri attributi presenti nello schema (`DB.costo_conn`, `Relazioni.costo_base` e `tuple_acc` dell'associazione tra `Query` e `Relazioni`).

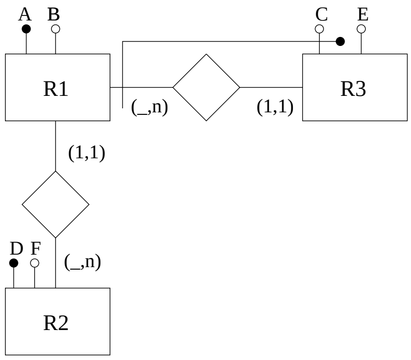
- Una chiave è in generale formata da più attributi. Poiché il modello E/R non fornisce primitive per identificare insiemi, risulta necessario identificare esplicitamente le `Chiavi` mediante l'attributo `KeyID`.

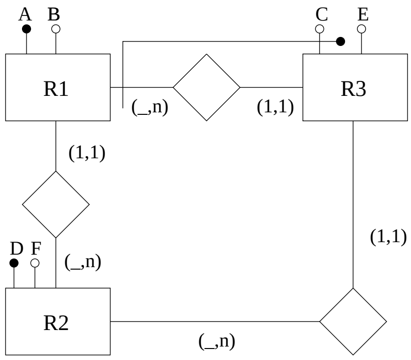
Esercizio 2

La copertura minima, $\mathcal{G}$, si ottiene da $\mathcal{F}$ eliminando la dipendenza ridondante $AC \rightarrow D$, ovvero:

$$\mathcal{G} = \{A \rightarrow BD, AC \rightarrow E, D \rightarrow F\}$$

Lo schema E/R di sinistra è conforme ai vincoli in $\mathcal{G}$, mentre quello di destra rappresenta i vincoli in $\mathcal{F}$. In entrambi le cardinalità non vincolate sono rappresentate con il simbolo '_'. L'associazione tra R_2 e R_3 è da considerarsi ridondante, in quanto il valore di D corrispondente ad una coppia di valori (A, C) dipende solo da A, e quindi è derivabile senza necessità di coinvolgere l'entità R_3.





Esercizio 3

Le istruzioni SQL che risolvono l'esercizio sono, rispettivamente:

```
Select F.Progetto
From   Afferenze A, Finanziamenti F
Where  A.Ricercatore = F.Ricercatore
and    A.Dipartimento = 'd'
and    F.Anno = 1996 ;

Select A.Ricercatore
From   Afferenze A
Where  A.Dipartimento = 'd'
and    A.Ricercatore in ( Select   F.Ricercatore
                          From     Finanziamenti F
                          Where    F.Anno <= 1996
                          Group by F.Ricercatore
                          Having   sum(F.Residuo) <= 0.2 * sum(F.Importo) );

Select   A.Dipartimento, sum(F.Importo)
From     Afferenze A, Finanziamenti F
Where    A.Ricercatore = F.Ricercatore
and      F.Ricercatore not in ( Select   F1.Ricercatore
                                From     Finanziamenti F1
                                Group by F1.Ricercatore
                                Having   sum(F.Importo) < tot_min )
Group by A.Dipartimento ;
```

Esercizio 4

La strategia di esecuzione ipotizzata si basa su un *nested-loops join* in cui si accede alla relazione interna `R` mediante un indice unclustered costruito su `R.B`.

Se tutti gli $ET_{\mathtt{S}}$ valori residui di `S.B` fossero presenti in `R`, allora la soluzione sarebbe data dalla formula:

$$ET_{\mathtt{S}} \times (NI_{\mathtt{R.B}} + \lceil NL_{\mathtt{R.B}}/NK_{\mathtt{R.B}} \rceil + \Phi(NT_{\mathtt{R}}/NK_{\mathtt{R.B}}, NP_{\mathtt{R}}))$$

La soluzione corretta considera che solo X valori sono presenti in `R`. Poiché solo per tali valori è necessario accedere alle pagine dati, il costo si esprime come:

$$ET_{\mathtt{S}} \times NI_{\mathtt{R.B}} + X \times (\lceil NL_{\mathtt{R.B}}/NK_{\mathtt{R.B}} \rceil + \Phi(NT_{\mathtt{R}}/NK_{\mathtt{R.B}}, NP_{\mathtt{R}})) + (ET_{\mathtt{S}} - X) \times 1$$

in cui l'ultimo termine quantifica il costo, da pagare per ognuno dei valori non presenti in `R`, di lettura della foglia dell'indice su cui la ricerca ha termine con insuccesso.

L'insieme degli X valori per cui è necessario accedere alle pagine dati di `R` è dato dall'intersezione di due sottoinsiemi, di cardinalità rispettivamente $NK_{\mathtt{R.B}}$ e $ET_{\mathtt{S}}$, dei valori presenti inizialmente in `S`. Ciò permette di stimare X come:

$$X = \frac{NK_{\mathtt{R.B}} \times ET_{\mathtt{S}}}{NT_{\mathtt{S}}}$$