

Cognome e nome: _____ Matricola: _____ Turno: _____

Esame di SQL

Rispondere su fogli separati rispetto alla parte di teoria.

Riportare sui fogli i seguenti dati: cognome, nome, matricola e turno di laboratorio.

Le seguenti relazioni definiscono una base di dati “PFWWrestling” per gestire le informazioni di una federazione di wrestling. Gli attributi sottolineati sono le chiavi primarie delle relazioni.

WRESTLER(RingName, Nome, Cognome, Sesso, Peso, StatoProvenienza)

CINTURA(Campione, Titolo, NumeroDiSuccessi)

STABLE(NomeStable, Capitano*)

COMPOSIZIONESTABLE(NomeStable, NomeComponente)

Vincoli di integrità referenziale:

CINTURA(Campione) referencia WRESTLER(RingName),

STABLE(Capitano) referencia WRESTLER(RingName),

COMPOSIZIONESTABLE(NomeStable) referencia STABLE(NomeStable),

COMPOSIZIONESTABLE(NomeComponente) referencia WRESTLER(RingName),

STABLE(NomeStable, Capitano) referencia COMPOSIZIONESTABLE(NomeStable, NomeComponente).

“Sesso” può assumere i valori “Uomo” e “Donna”. Capitano può essere NULL. Titolo può assumere i valori “assoluto”, “intercontinentale” e “femminile”. Una stable è composta da almeno tre wrestler. Solo le donne possono avere un titolo femminile. I rimanenti attributi sono autoesplicativi.

Con riferimento alla base di dati "PFWWrestling" esprimere in SQL le seguenti interrogazioni.

Domanda 1 (bassa complessità).

Trovare gli stati di provenienza dei capitani di stable che hanno vinto almeno 3 volte il titolo assoluto.

Soluzione 1.

```
SELECT DISTINCT StatoProvenienza
FROM wrestler JOIN stable ON (Capitano = RingName)
      JOIN cintura ON (Capitano = Campione)
WHERE Titolo = 'assoluto' AND NumeroDiSuccessi >= 3;
```

Domanda 2 (media complessità).

Trovare ringname, nome e cognome dei wrestlers canadesi che fanno parte di almeno 5 stable i cui capitani hanno vinto il titolo intercontinentale.

Soluzione 2.

```
SELECT RingName, Nome, Cognome
FROM wrestler JOIN composizionestable c ON (RingName = NomeComponente)
      JOIN stable s ON (s.NomeStable = c.NomeStable)
      JOIN cintura ON (Capitano = Campione)
WHERE StatoProvenienza = 'Canada' AND Titolo = 'intercontinentale'
GROUP BY RingName, Nome, Cognome
HAVING COUNT(DISTINCT NomeStable)>=5;
```

Domanda 3 (alta complessità).

Tra le stable composte da soli uomini, trovare quelle che hanno il minimo numero di componenti.

Soluzione 3.

```
WITH StableSenzaDonne AS (  
    SELECT s.NomeStable, COUNT(s.NomeComponente) NumeroComponenti  
    FROM composizionestable s  
    WHERE NOT EXISTS (  
        SELECT *  
        FROM composizionestable c  
        JOIN wrestler w ON (w.RingName = c.NomeComponente)  
        WHERE c.NomeStable=s.NomeStable AND w.Sesso = 'Donna')  
    GROUP BY s.NomeStable)  
SELECT NomeStable  
FROM StableSenzaDonne  
WHERE NumeroComponenti = (SELECT MIN(NumeroComponenti)  
                           FROM StableSenzaDonne);
```

Esame di Teoria (rispondere su fogli separati rispetto alla parte di SQL)

Domanda 1 (9 punti).

Con riferimento alla base di dati “PFWWrestling”:

- A. Esprimere in algebra relazionale l'interrogazione
Elencare i wrestler uomini che hanno peso massimo.
- B. Esprimere, nel calcolo dei predicati su tuple con dichiarazione di range, la seguente domanda:
Elencare le stable composte da soli campioni e capitanate da wrestler canadesi.

Soluzione 1.

A. Una possibile soluzione è la seguente:

$$\pi_{W1.RingName} \left(\rho_{W1 \leftarrow WRESTLER} \left(\sigma_{sesso='Uomo'}(wrestler) \right) \right) \\ - \pi_{W1.RingName} \left(\rho_{W1 \leftarrow WRESTLER} \left(\sigma_{sesso='Uomo'}(wrestler) \right) \right) \bowtie_{W1.Peso < W2.Peso} \\ \rho_{W2 \leftarrow WRESTLER} \left(\sigma_{sesso='Uomo'}(wrestler) \right)$$

B. Una possibile soluzione è la seguente:

$$\{s.NomeStable \mid s(STABLE) \mid \exists w(WRESTLER)(s.Capitano=w.RingName \wedge w.StatoProvenienza='Canada' \wedge \\ \forall cs(COMPOSIZIONESTABLE)(cs.NomeStable=s.NomeStable \Rightarrow \\ \exists c(CINTURA)(cs.NomeComponente=c.Campione)))\}$$

Domanda 2 (8 punti).

- A. Riportare la definizione di BCNF.
- B. Data una relazione $R(A,B,C,D,E,F)$ e l'insieme di dipendenze funzionali $F = \{AB \rightarrow EF, A \rightarrow CD, C \rightarrow B, E \rightarrow F\}$, dire se F è in 3NF motivando la risposta e se non lo è decomporla in 3NF mostrando tutti i passaggi.
Dire se il risultato è in BCNF motivando la risposta.

Soluzione 2.

A. Vedere libro di testo o slide.

B. Per dire se R è in 3NF, trovo prima tutte le chiavi della relazione R . Dato che A non compare mai nei conseguenti, deve appartenere a ogni chiave. Possiamo vedere che $A+ = ABCDEF$, cioè tutti gli attributi di R , quindi A è superchiave. Dato che è una superchiave composta da un solo attributo, è già minimale, quindi è anche chiave. Non ci sono altre chiavi perché, come abbiamo osservato, dovrebbero contenere A , quindi non sarebbero superchiavi non minimali.

R non è in 3NF perché, ad es., $E \rightarrow F$ non è in 3NF perché non è banale, E non è superchiave e F non è primo.

Trasformo le d.f. mettendo un solo attributo a destra: $F' = \{AB \rightarrow E, AB \rightarrow F, A \rightarrow C, A \rightarrow D, C \rightarrow B, E \rightarrow F\}$.

Calcolo un insieme di copertura minimale.

Elimino gli attributi estranei. Dato che A è chiave, sicuramente in $AB \rightarrow E$ e in $AB \rightarrow F$ B è estraneo perché posso ricavare E e F con il solo attributo A . Le altre d.f. hanno un solo attributo a sinistra, quindi non hanno attributi estranei. Ottengo $F'' = \{A \rightarrow E, A \rightarrow F, A \rightarrow C, A \rightarrow D, C \rightarrow B, E \rightarrow F\}$.

Elimino le dipendenze ridondanti.

$A \rightarrow E$ è ridondante? Dato $F^* = \{A \rightarrow F, A \rightarrow C, A \rightarrow D, C \rightarrow B, E \rightarrow F\}$, controllo se $E \subseteq AF^*+$. Dato che E non compare nel conseguente di nessuna d.f. in F^* , E non sarà nella chiusura di A e quindi $A \rightarrow E$ non è ridondante. $A \rightarrow F$ è ridondante? Dato $F^* = \{A \rightarrow E, A \rightarrow C, A \rightarrow D, C \rightarrow B, E \rightarrow F\}$, controllo se $F \subseteq AF^*+$. Ricavo F da A con $A \rightarrow E$ e $E \rightarrow F$, quindi $A \rightarrow F$ è ridondante e la elimino.

$A \rightarrow C$ è ridondante? Dato $F^* = \{A \rightarrow E, A \rightarrow D, C \rightarrow B, E \rightarrow F\}$, controllo se $C \subseteq AF^*+$. Dato che C non compare nel conseguente di nessuna d.f. in F^* , sicuramente C non è contenuta nella chiusura di A, quindi $A \rightarrow C$ non è ridondante.

$A \rightarrow D$ è ridondante? Dato $F^* = \{A \rightarrow E, A \rightarrow C, C \rightarrow B, E \rightarrow F\}$, controllo se $D \subseteq AF^*+$. Dato che D non compare nel conseguente di nessuna d.f. in F^* , sicuramente D non è contenuto nella chiusura di A, quindi $A \rightarrow D$ non è ridondante.

$C \rightarrow B$ è ridondante? Dato $F^* = \{A \rightarrow E, A \rightarrow C, A \rightarrow D, E \rightarrow F\}$, controllo se $B \subseteq CF^*+$. Dato che C non compare nell'antecedente di nessuna d.f. in F^* , la chiusura contiene solo C stesso e non B, quindi $C \rightarrow B$ non è ridondante.

$E \rightarrow F$ è ridondante? Dato $F^* = \{A \rightarrow E, A \rightarrow C, A \rightarrow D, C \rightarrow B\}$, controllo se $F \subseteq EF^*+$. Dato che E non compare nell'antecedente di nessuna d.f. in F^* , la chiusura contiene solo E stesso e non F, quindi $E \rightarrow F$ non è ridondante. La copertura minimale è $\{A \rightarrow E, A \rightarrow C, A \rightarrow D, C \rightarrow B, E \rightarrow F\}$.

Decompongo in 3NF. Raggruppando per antecedente, ottengo: $R_1(A,C,D,E)$, $R_2(C,B)$, $R_3(E,F)$.

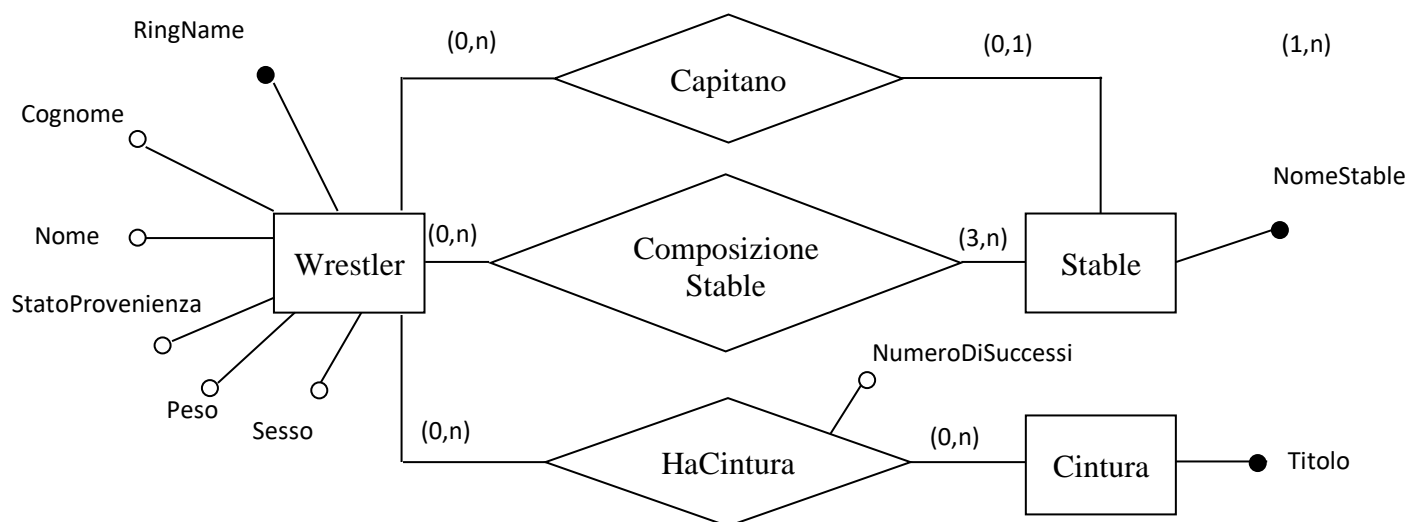
Nessuna relazione è sottoinsieme di un'altra. Inoltre R_1 contiene una chiave di R, quindi ho terminato la normalizzazione in 3NF.

Il risultato è in BCNF perché ogni d.f. ha come antecedente una chiave.

Domanda 3 (8 punti).

Progettare uno schema concettuale che rappresenti la base di dati "PWWrestling". Ricordarsi di riportare, sotto forma di regole aziendali, tutti i vincoli non esprimibili attraverso l'ER.

Soluzione 3.



Regole aziendali:

- $Wrestler.Sesso \in \{“Uomo”, “Donna”\}$
- $Cintura.Titolo \in \{“assoluto”, “intercontinentale”, “femminile”\}$
- Se un Wrestler partecipa all’associazione HaCintura con $Cintura.Titolo=“femminile”$ allora $Wrestler.Sesso=“Donna”$.
- Le occorrenze di Capitano sono un sottoinsieme delle occorrenze di ComposizioneStable (cioè un capitano di una stable deve fare parte della stable)

Domanda 4 (8 punti).

Con riferimento alla base di dati “PWF Wrestling”, si considerino due indici IW e IS realizzati con B+-tree rispettivamente su wrestler(RingName) e stable(Capitano) e i seguenti dati quantitativi:

$N_{foglie}(IW) = 25$

$CARD(wrestler) = 2000$

$N_{page}(wrestler) = 100$

$N_{foglie}(IS) = 10$

$CARD(stable) = 6000$

$N_{page}(stable) = 15$

$VAL(Capitano, stable) = 500$

Confrontare le stime dei costi della seguente interrogazione, già logicamente ottimizzata, eseguita sfruttando il join con nested loop con indice e con block-nested loop con $B=6$ (fornire sia le formule risolutive che i risultati numerici). Si tenga presente che un record di wrestler occupa circa 200 Byte.

$(\sigma_{RingName='John\ Cena'(wrestler)}) \bowtie_{RingName=Capitano} stable$

Soluzione 4.

Il costo della selezione è dato da: $C(\sigma) = CI + CD$

$CI = \lceil fp \cdot N_{foglie}(IW) \rceil = \lceil 1/CARD(wrestler) \cdot N_{foglie}(IW) \rceil = \lceil 1/2000 \cdot 25 \rceil = 1$

$CD = \min\{Erec, N_{page}(wrestler)\} = \min\{\lceil fp \cdot CARD(wrestler) \rceil, N_{page}(wrestler)\}$

$CD = \min\{\lceil 1/2000 \cdot 2000 \rceil, 100\} = \min\{1, 100\} = 1$

$C(\sigma) = CI + CD = 1 + 1 = 2$ accessi

Il costo del join (nested loop con indice) è dato da: $C(\bowtie) = N_{page}(\sigma) + CARD(\sigma) \cdot (CI + CD)$

$CARD(\sigma) = 1$ record perché selezioniamo usando la chiave primaria

$N_{page}(\sigma) = 1$ (una pagina occupa 4 KB / 8 KB, un record di wrestler occupa 200 B quindi è sufficiente una pagina per contenere il record).

Calcoliamo ora il costo del join con nested loop con indice:

$CI = \lceil fp \cdot N_{foglie}(IS) \rceil = \lceil 1/VAL(Capitano, stable) \cdot N_{foglie}(IS) \rceil = \lceil 1/500 \cdot 10 \rceil = 1$

$CD = \min\{Erec, N_{page}(stable)\} = \min\{\lceil fp \cdot CARD(stable) \rceil, N_{page}(stable)\}$

$CD = \min\{\lceil 1/500 \cdot 6000 \rceil, 15\} = \min\{12, 15\} = 12$

Quindi: $C(\bowtie) = 1 + 1 \cdot (1 + 12) = 14$

Il costo totale con nested loop con indice è quindi dato da: $C_{tot} = C(\sigma) + C(\bowtie) = 2 + 14 = 16$ accessi

Il costo del block-nested loop è dato da

$$C(\bowtie) = \text{Npage}(\sigma) + \lceil \text{Npage}(\sigma) / (B - 1) \rceil \cdot (\text{Npage}(\text{stable}))$$

$$C(\bowtie) = 1 + \lceil 1/5 \rceil \cdot 15 = 16$$

Il costo totale con block-nested loop è quindi $C_{\text{tot}} = C(\sigma) + C(\bowtie) = 2 + 16 = 18$ accessi

Il costo del join con nested loop con indice è inferiore al costo del join con nested loop con block-nested loop. L'ottimizzatore userà quindi il join con nested loop con indice.