

Matricola: _____

Cognome: _____

Nome: _____

Basi di dati

Il Prova in itinere del 27 febbraio 2018

Durata 2h

Avvertenze: è severamente vietato consultare libri e appunti.

Domande di teoria

- a) (2) Si illustri l'operatore di join naturale dell'algebra relazionale (sintassi, semantica, esempio d'uso)

- b) (2) Date le due seguenti relazioni: $R1(\underline{A}, B, C, D)$ e $R2(\underline{B}, E, F^*, G)$ si scriva in algebra relazionale:

b.1) un'espressione che restituisca una relazione con un solo attributo K contenente l'insieme dei valori distinti contenuti nell'attributo C di R1 che sono presenti anche nell'attributo G di R2;

b.2) un'espressione ottimizzata che contenga **solo** un join naturale, una selezione su R1 e una proiezione su R2 e produca come risultato le tuple t di R1 tali che $t[D] < 5$ e $t[C] = 9$ e per le quali esista una tupla t' di R2 dove $t[B] = t'[B]$.

ALGEBRA RELAZIONALE (è obbligatorio rispondere ai quesiti 1.a, 1.b, 1.c e 4.a)

Dato il seguente schema relazionale contenente le informazioni che descrivono una serie di tornei di tennis svolti o programmati nel 2018:

TORNEO(Nome, Città, Stato);

GIORNATA(Torneo, NumeroGiornata, Data)

TENNISTA(Codice, Nome, Cognome, Età, Nazionalità)

PARTITA(Torneo, NumeroGiornata, OraInizio, Giocatore1, Giocatore2, Durata*, Vincitore*)

Vincoli d'integrità referenziale:

GIORNATA.Torneo → TORNEO

PARTITA.(Torneo, NumeroGiornata) → GIORNATA, PARTITA.Giocatore1 → TENNISTA,

PARTITA.Giocatore2 → TENNISTA, PARTITA.Vincitore → TENNISTA

Ulteriori vincoli d'integrità: NumeroGiornata ∈ {1,2,3,4,5,6,7,8,9,10,11,12,13,14,15}

1. Formulare in algebra relazionale **ottimizzata** le seguenti interrogazioni:

1.a (3) *Trovare il nome e la città dei tornei spagnoli dove nella prima e nella seconda giornata non sono previste partite dopo le 21.30.*

1.b (3) *Trovare il nome e il cognome dei tennisti italiani che hanno giocato almeno due partite nel torneo "T3".*

1.c (3) *Trovare le partite che si svolgono oggi o domani dopo le 11.30 nei tornei che non si tengono in Francia riportando il nome del torneo, la città dove si svolge, la data e l'ora d'inizio della partita e il cognome e la nazionalità dei due giocatori.*

2. Formulare in algebra relazionale le seguenti interrogazioni:

2.a (3) *Trovare la partita di durata minima che si è svolta in Italia riportando il nome del torneo, la città dove si è svolto, la data della partita e il cognome del vincitore.*

2.b (2) *Trovare il nome e il cognome dei tennisti che hanno partecipato (sono coinvolti in almeno una partita) a tutti i tornei svoltisi in USA.*

3. Supponendo che le relazioni abbiano le seguenti cardinalità:

- **TORNEO:** 50 (tornei italiani = 5)
- **GIORNATA:** 250
- **TENNISTA:** 90 (tennisti francesi = 10)
- **PARTITA:** 1500 (partite vinte da tennisti francesi in tornei italiani = 50)

3.a (3) ipotizzando che per tutte le partite si sia registrato il vincitore, calcolare la dimensione dei risultati intermedi (in termini di numero di valori) in tutti i nodi dell'albero che rappresenta la seguente interrogazione:

$$\Pi_{\{\text{Cognome, Nome, Città}\}}(\sigma_{\text{Nazionalità}='Francia' \wedge \text{Stato}='Italia'}((\text{PARTITA} \bowtie_{\text{Torneo} = \text{Nome}} \text{TORNEO}) \bowtie_{\text{Vincitore} = \text{Codice}} \Pi_{\{\text{Codice, Cognome, Nazionalità}\}}(\text{TENNISTA})))$$

3.b (3) produrre la versione ottimizzata della precedente interrogazione

3.c (3) calcolare la dimensione dei risultati intermedi (in termini di numero di valori) in tutti i nodi dell'albero che rappresenta la versione ottimizzata prodotta al punto precedente.

SQL (sono ammesse solo le parole chiave: SELECT, FROM, WHERE, EXISTS, IN, AND, OR, NOT, DISTINCT)

4. Dato lo schema relazionale sopra riportato, formulare in SQL le seguenti interrogazioni:

4.a (4) *Trovare per ogni giornata del torneo di nome "T1" la prima partita della giornata, riportando nel risultato la giornata, l'ora di inizio e il cognome dei giocatori della prima partita.*