



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 30 Giugno 2021

Regole dell'esame:

- Non è possibile utilizzare alcun materiale, né appunti, né il libro.
- **Inserire le risposte nei riquadri che seguono i testi degli esercizi**
- L'uso della matita per mostrare la soluzione degli esercizi è fatta a rischio e pericolo degli studenti. Nessuna rivendicazione verrà accettata in caso parte della soluzione si cancelli, per via delle caratteristiche transitorie dei tratti a matita.
- Il massimo possibile è 32 punti
- Occorre avere acquisito almeno metà dei punti dell'Esercizio 3 (cioè 3.5 punti).
- **Coloro che siano sorpresi a copiare o a far copiare:**
 - Dovranno lasciare l'aula dell'esame
 - Riceveranno 0 punti "di ufficio"
 - Dovranno saltare l'appello successivo

Nome: _____ Num. Matric.: _____

Esercizio 1: Diagramma ER (8 punti)

*Si richiede di produrre lo schema concettuale Entità-Relazione di un'applicazione relativa ai concorsi banditi da istituzioni pubbliche. **Disegnare il diagramma ER nel riquadro della pagina che segue.***

Il sistema è condiviso tra diverse istituzioni, di ognuna delle quali è di interesse sapere il nome dell'istituzione, la via e il codice ISTAT del comune in cui ha la sede principale.

Per ogni istituzione, si vuole sapere le persone che lavorano presso tale istituzione ed i bandi di concorso che ha eventualmente emesso.

Di ogni persona interessa il codice fiscale (identificativo), il nome, il cognome, la data di nascita. Una persona può lavorare presso una istituzione (ma anche essere disoccupata). Alcune persone possono anche fungere da commissari giudicatori di concorso: per loro, si vuole sapere quale sia il titolo di studio.

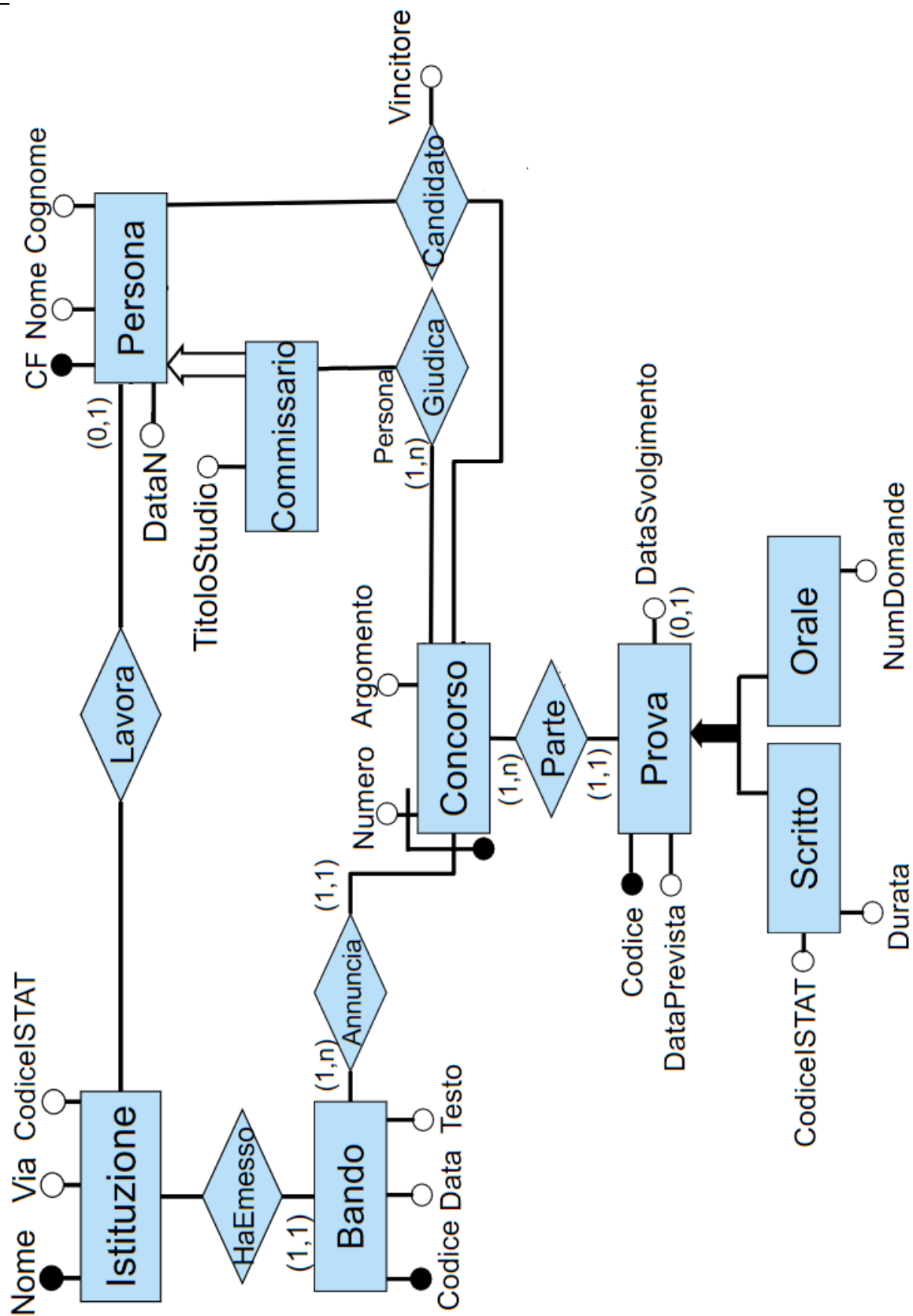
Di ogni bando interessa il codice identificativo, la data di pubblicazione, il testo ed i concorsi (almeno uno) che sono annunciati nel bando stesso.

Di ogni concorso interessa il bando in cui è annunciato, il numero (unico nell'ambito del bando in cui è annunciato ma diversi bandi possono avere lo stesso numero concorso), l'argomento del concorso (giurisprudenza, ingegneria civile, amministrazione, ecc.) e le prove di cui è composto (almeno una), identificata da un codice identificativo. Per ogni prova, si vuole conoscere la data di svolgimento prevista e successivamente la data di svolgimento effettiva, quando questa avviene.

Le prove sono o scritte o orali e di ogni prova scritta interessa anche la durata e il codice ISTAT del comune dove si prevede si svolga, mentre di ogni prova orale interessa anche il numero di domande previste.

Di ogni concorso interessano anche le persone (almeno una) che hanno partecipato come commissari, le persone che hanno partecipato come candidati (se ci sono) e, tra queste ultime, quelle che sono risultate vincitrici (se ci sono).

Soluzione Esercizio 1



Le cardinalità (0,n) omesse per leggibilità

Nome: _____ Num. Matric.: _____

Esercizio 2: Progettazione Logica (7 punti)

A partire dal Diagramma ER dell'Esercizio 1, produrre uno schema relazionale del database nel riquadro sottostante. Indicare i vincoli di chiave e gli attributi che ammettono valori nulli. **Occorre ottimizzare la seguente operazione: quando si accede alla durata di una prova o al numero di domande, si vuole sempre conoscere il concorso di cui fa parte** Illustrare come ristrutturare l'ER per essere traducibile in uno schema relazionale.¹

Soluzione Esercizio 2 (Parziale)

¹ Allo scopo di mostrare la ristrutturazione del diagramma ER, è possibile semplicemente mostrare i cambiamenti apportati al diagramma nel riquadro soluzione dell'Esercizio 1, utilizzando una penna di diverso colore (non rossa!)

Esercizio 3: Algebra Relazionale & SQL (7 punti)

Si consideri la seguente base di dati dei diversi servizi sociali in città diverse di Italia:

- **SERVIZI_SOCIALI** (Citta, Servizio, Anno, Spesa)
- **POSIZIONE** (Citta, Regione, Abitanti)

dove non è possibile avere due città con lo stesso nome. Una tupla (Padova, Babysitting, 2019, 30000) in SERVIZI SOCIALI indica che Padova ha speso nel 2019 la cifra di 30000€ per il servizio sociale Babysitting.

- A. Restituire le città che forniscono esattamente un servizio sociale (2 punti).²

S1=SERVIZI_SOCIALI

S2=SERVIZI_SOCIALI

CITTA_DUE =

$\pi_{S1.CITTA} (S1 \bowtie_{S1.CITTA=S2.CITTA \text{ AND } S1.SERVIZIO \neq S2.SERVIZIO} S2)$

$\pi_{CITTA} (S1) \setminus CITTA_DUE$

² Si assuma che l'operatore di join $A \bowtie B$ **senza condizioni** mantenga le tuple di $A \times B$ con valori uguali su attributi uguali (join naturale). Se **una condizione C è specificata**, \bowtie_C mantiene le tuple di $A \times B$ per cui la condizione C è vera.

Nome: _____ Num. Matric.: _____

B. Scrivere una query in Standard SQL che restituisce la regione in cui c'è la città che ha speso di più in servizi sociali nel 2020. (2.5 punti).

```
CREATE VIEW CITTA-SPESA
(CITTA, SOMMA) AS
SELECT CITTA, SUM(SPESA)
FROM SERVIZI_SOCIALI S
WHERE CITTA
GROUP BY CITTA;

SELECT DISTINCT REGIONE
FROM CITTA-SPESA C, POSIZIONE P
WHERE C.CITTA=P.CITTA AND SOMMA =
      (SELECT MAX(SOMMA) FROM REGIONE-CITTA-SPESA)
```

C. Nel riquadro, scrivere una query in Standard SQL che, per ogni città con almeno 200000 abitanti con una spesa media di servizio maggiore di 50000 €, restituisce la città con il numero di servizi diversi forniti (2.5 punti)

```
SELECT CITTA, COUNT(DISTINCT SERVIZIO)
FROM SERVIZI_SOCIALI S, POSIZIONE P
WHERE S.CITTA=P.CITTA AND ABITANTI>=200000
GROUP BY CITTA
HAVING AVG(SPESA)>50000
```

Esercizio 4: Normalizzazione (5 punti)

Sia data la relazione $R(A,B,C,D,E,F)$ e l'insieme di dipendenze associato $F=\{C\rightarrow CABDEF, AD\rightarrow E, BE\rightarrow C, B\rightarrow A\}$:

- Trovare la/e chiave/i di R , motivando la risposta.
- Indicare quali dipendenze violano la 3NF, motivando la risposta
- Calcolare la copertura ridotta di F . Aggiungere solo la soluzione.
- Effettuare una decomposizione in 3NF.

Punto a

Si considerano le seguenti chiusure

$C^+=\{C,A,B,D,E,F\}$

$AD^+=\{A,D,E\}$

$BE^+=\{C,A,B,D,E,F\}$

$B^+=\{B,A\}$

Quindi, C e BE sono chiavi della relazione.

Punto b

$B\rightarrow A$ viola la 3NF perché B non è superchiave e A non appartiene a nessuna delle chiavi.

Punto c

La copertura ridotta è $G=\{C\rightarrow B, C\rightarrow D, C\rightarrow F, BE\rightarrow C, B\rightarrow A, AD\rightarrow E\}$

Punto d

1. G è partizionato in sottoinsiemi tali che dip. funz. $X \rightarrow A$ e $Y \rightarrow B$ sono insieme se $X^+ = Y^+$

$\{C\rightarrow B, C\rightarrow D, C\rightarrow F, BE\rightarrow C\}, \{B\rightarrow A\}, \{AD\rightarrow E\}$

2. Viene costruita una relazione per ogni sottoinsieme:

$R_1(C,B,D,F,E), R_2(B,A), R_3(AD,E)$

3. Se esistono due relazioni $A(X)$ and $T(Y)$ con $X \subseteq Y$, A viene eliminata:

Non accade

4. Se esiste una chiave K per quale non esiste una relazione che contiene tutti gli attributi di K , viene aggiunta una relazione $T(K)$:

Non accade

5. Indicare le chiavi delle relazioni ottenute dalla normalizzazione

$R_1(C,B,D,F,E)$ con chiave C (aggiungendo anche chiave BE , si ottiene una decomposizione BCNF).

$R_2(B,A)$ con chiave B

$R_3(A,D,E)$ con chiave AD

Nome: _____ Num. Matric.: _____

Esercizio 5: Quiz (5 punti)

Rispondere alle seguenti domande, sottolineando quale risposta è corretta (solo una è corretta).

Domanda 1 (1.5 Punti)

Data la query `SELECT * FROM S WHERE Z>4 AND X=5` sulla relazione $S(X, Y, Z, W)$. Quale dei seguenti indici in genere assicura le migliori performance in termini di velocità dell'esecuzione della query?

1. Indice B+Tree sulla coppia (X,Z)
2. Indice Hash sulla coppia (Z,X)
3. Indice Hash sulla coppia (X,Z)
4. Indice B+Tree sulla coppia (Z,X)

Domanda 2 (1.5 Punti)

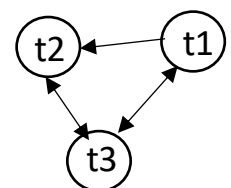
Date una qualsiasi istanza della relazioni $R(\underline{A}, B, C)$ e $S(\underline{D}, E, F)$ dove (1) le uniche chiavi di R e S sono rispettivamente $\{A\}$ e $\{D\}$, (2) non esistono chiavi esterne. Indicato con $|X|$ il numero di tuple di una relazione X, quale è vera tra le seguenti affermazioni relative al numero di tuple in $R \bowtie_{A=E} S$?

1. $0 \leq |R \bowtie_{A=E} S| \leq |R|$
2. $0 \leq |R \bowtie_{A=E} S| \leq |R| * |S|$
3. $|R \bowtie_{A=E} S| = |S|$
4. $0 \leq |R \bowtie_{A=E} S| \leq \max(|R|, |S|)$

Domanda 3 (2 Punti)

Dato il seguente schedule S, con grafo di conflitti in figura:

$r1(a) \ r1(b) \ w3(c) \ r2(c) \ r2(a) \ w2(a) \ r3(a) \ r1(c) \ w3(c)$



Quale delle seguenti affermazioni è vera?

1. S è conflict-serializzabile ma **non** view-serializzabile
2. S **non** è conflict-serializzabile ma è view-serializzabile
3. S è **sia** conflict-serializzabile **che** view-serializzabile
4. S non è né conflict-serializzabile né view-serializzabile