



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Esempio di Esame

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 ad una ditta che estrae marmo per abitazioni da cave. Disegnare il diagramma ER nel riquadro della pagina che segue.

Di ogni blocco di marmo estratto interessa il codice (identificativo), l'anno di estrazione, il peso e la cava da cui è stato estratto.

Per il dominio di estrazione del marmo, è interesse memorizzare i dati di certi luoghi geografici. Ogni luogo geografico ha un codice identificativo e l'area che occupa in chilometri quadrati.

Le cave di interesse sono quelle dalle quali è stato estratto almeno un blocco. Ogni cava è un luogo geografico di cui interessa anche conoscere l'altitudine e la regione (esattamente una) in cui è situato il suo territorio.

Da ogni blocco di marmo si producono almeno una lastra di marmo e di ogni lastra interessa il blocco da cui è stato prodotto (uno ed uno solo), la superficie, il numero (unico nell'ambito del blocco di marmo da cui è stato prodotto), l'eventuale abitazione in cui viene usata.

Ogni abitazione che interessa all'applicazione usa almeno una lastra di marmo e di ognuna di tali abitazioni interessa il codice (identificativo), la classe e le eventuali banche che hanno partecipato all'asta per quell'abitazione, con la data di partecipazione della banca all'asta.

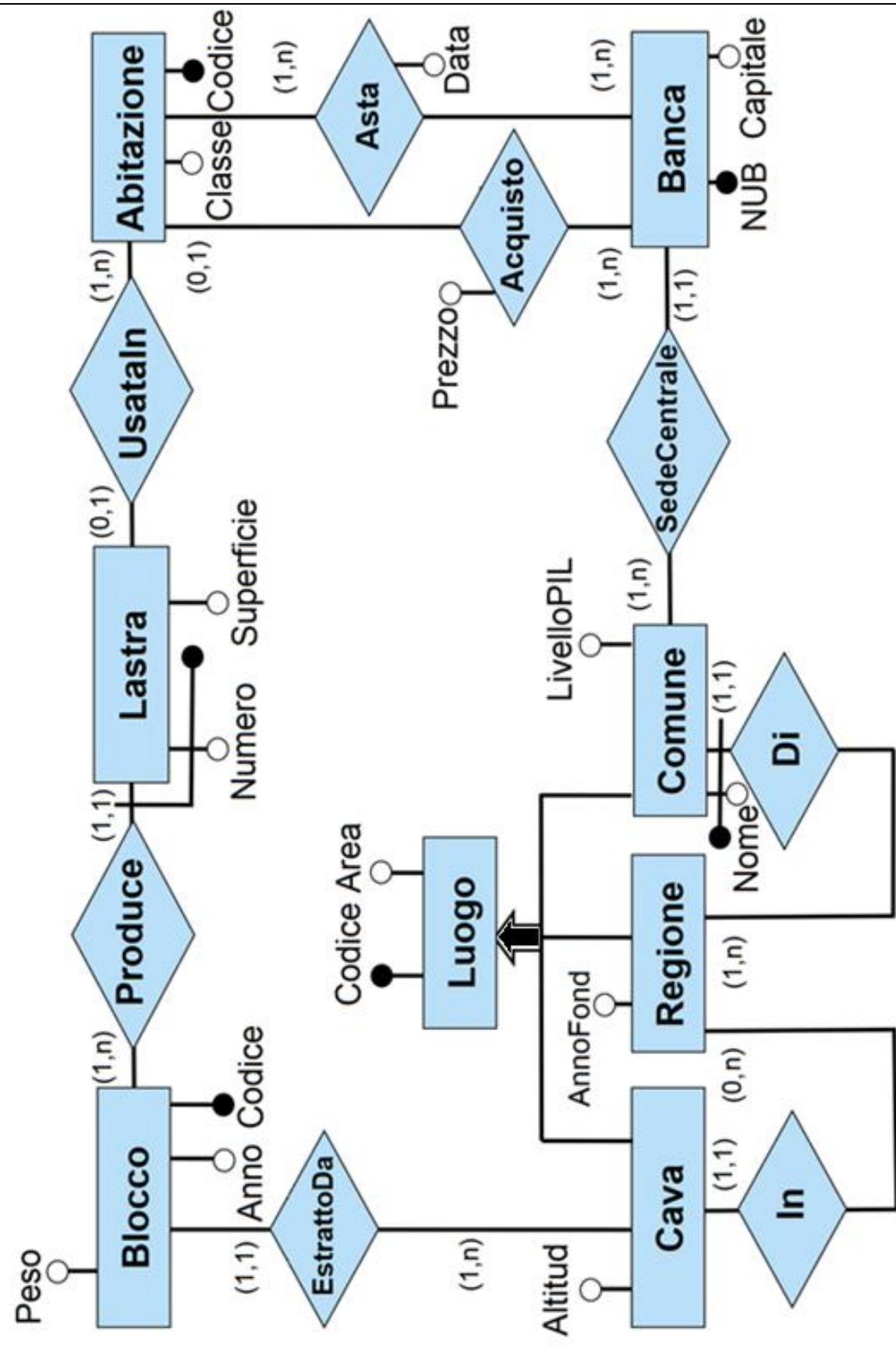
Tra le banche che hanno partecipato all'asta per una certa abitazione, dopo tale asta, è di interesse sapere la banca, se esiste, che ha acquistato l'abitazione stessa, con il relativo prezzo di acquisto.

Di ogni banca interessa il “codice unico bancario” (identificativo), il capitale sociale ed il comune in cui si trova la sede centrale.

Di ogni comune, oltre alle proprietà di tutti i luoghi geografici, interessa il livello del PIL, la regione in cui si trova ed il nome (unico nell'ambito della regione in cui si trova).

Di ogni regione, oltre alle proprietà di tutti i luoghi geografici, interessa l'anno della sua fondazione.

Soluzione Esercizio 1



Nome: _____ Num. Matric.: _____

Esercizio 2: Progettazione Concettuale (7 punti)

A partire dallo ER concettuale al punto 1, produrre uno schema relazionale del database nel riquadro sottostante. Indicare la chiave primaria, i vincoli di chiave esterne, quali attributi ammettono valori nulli, ed ogni altro vincolo rilevante. Mostrare il diagramma ER ristrutturato per eliminare i costrutti non direttamente rappresentabili.¹

Soluzione Esercizio 2

L'entità **Luogo** viene eliminata, insieme alla generalizzazione. Gli attributi **Codice** e **Area** vengono replicati in **Cava**, **Regione** e **Comune**.

Tutti gli attributi seguiti da un asterisco ammettono valori nulli (Nota: all'esame è possibile rappresentare i valori nulli in qualsiasi modo perché non c'è uno standard. L'importante è che sia chiaro quale vengono rappresentati come nulli)

Regione(Codice, Area, AnnoFondazione)

Cava(Codice, Area, Altitudine, CodRegione)

- Foreign key: Cava.CodRegione → Regione.Codice

Comune(Codice, CodRegione, Nome, Area, LivelloPIL)

- Foreign key: Comune.CodRegione → Regione.Codice
- (CodRegione, Nome) è una seconda chiave.

Banca(NUB, Capitale, CodRegSedeCentrale, NomeComuneSedeCentrale)

- Foreign key:
Banca.(CodRegSedeCentrale, NomeComuneSedeCentrale) →
Comune(CodRegione, Nome)

Abitazione(Codice, Classe, BancaAcquisto*, PrezzoAcquisto*)

- Foreign key: Abitazione.BancaAcquisto → Banca.NUB

Asta(NUB-Banca,CodiceAbitazione, Data)

- Foreign key: Asta. NUB-Banca → Banca.NUB
- Foreign key: Asta.CodiceAbitazione → Abitazione.Codice

Blocco(Codice, Anno, Peso, CodiceCava)

- Foreign key: Blocco.CodiceCava → Cava.Codice

Lastra(Numer,Blocco,Superficie, AbitazioneUso*)

- Foreign key: Lastra.Blocco → Blocco.Codice
- Foreign key: Lastra.Abitazione → Abitazione.Codice

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 per la registrazione dei concorsi, i candidati e gli esiti:

- CANDIDATO(CF, Nome, Cognome)
- PARTECIPA(CF, CodConcorso, Esito)
- CONCORSO(CodConcorso, Descrizione, Anno)

dove Esito può essere ‘positivo’ o ‘negativo’ (usare queste due costanti).

- A. Nel riquadro, scrivere una Query in Algebra Relazione che restituisce i nomi e cognomi di tutti i candidati con esito negativo per almeno un concorso del 2019 (2 punti).²

```
 $\Pi_{\text{Nome}, \text{Cognome}}(\text{CANDIDATO} \bowtie (\sigma_{\text{Esito}='\text{negativo}' \text{ AND } \text{Anno}=2019}(\text{CONCORSO} \bowtie \text{PARTECIPA})))$ 
```

- B. Nel riquadro, scrivere un’interrogazione SQL che restituisca tutti i codici dei concorsi che hanno avuto solo esiti negativi, senza duplicati (2.5 punti).

```
SELECT CodConcorso FROM CONCORSO  
EXCEPT  
SELECT CodConcorso FROM PARTECIPA WHERE  
Esito='Positivo'
```

oppure

```
SELECT CodConcorso FROM CONCORSO  
WHERE CodConcorso NOT IN  
(SELECT CodConcorso FROM PARTECIPA  
WHERE Esito='Positivo')
```

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

Nome: _____ Num. Matric.: _____

C. Nel Riquadro, scrivere un'interrogazione SQL che restituisca il codice fiscale dei candidati che hanno partecipato a TUTTI i concorsi con descrizione “Banca di Italia” a partire dal 2000 (2.5 punti)

La seguente è una soluzione ma non l'unica possibile:

```
SELECT CF
FROM PARTECIPA P, CONCORSO C
WHERE P.CodConcorso = C CodConcorso AND
Descrizione='Banca di Italia' AND Anno>2000
GROUP BY CF
HAVING COUNT(*) =
    (SELECT COUNT(*) FROM CONCORSO
     WHERE Descrizione='Banca di Italia'
     AND Anno>2000)
```

Esercizio 4: Transazioni (5 punti)

Indicare e motivare se lo schedule è conflict-serializzabile

$r_4(y)w_1(z)r_2(y)w_3(x)w_1(y)r_1(x)r_3(z)w_5(z)w_5(y)w_4(z)r_4(x)$

In caso sia conflict-serializzabile, indicare come le transazioni possono essere riordinate per ottenere uno schedule seriale conflict-equivalente.

Il grafo dei conflitti ha un ciclo dalla transazione 4 alla transazione 1, dovuta a $r_4(y) w_1(z) \dots w_1(y) w_4(z)$.

Quindi, lo schedule non è conflict-serializzabile.

Esercizio 5: Quiz (5 punti)

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

Domanda 1 (1.5 Punti)

Data la relazione e $R(\underline{A}, B, C)$ la query $\text{SELECT } * \text{ FROM } R \text{ ORDER BY } C$, quale dei seguenti indici velocizza l'esecuzione della query?

- (1) Indice Hash su C
- (2) Indice B+Tree su C
- (3) Indice Hash su A, B, C,
- (4) Indice B+Tree su A, B, C

Domanda 2 (1.5 Punti)

Date due relazioni $R(\underline{A}, B, C)$ e $S(\underline{D}, E, F)$ dove (1) le uniche chiavi di R e S sono quelle primarie, (2) C è chiave esterna a D e (3) C non può essere NULL. Indicato con $|X|$ il numero di tuple di una relazione X, il join $R \bowtie_{C=D} S$ restituisce il seguente numero di tuple:

- (1) $|R|$
- (2) $|S|$
- (3) $|R|^*|S|$
- (4) Il minimo tra $|R|$ e $|S|$

Domanda 3 (2 Punti)

Data la relazioni $R(\underline{A}, B, C)$ con chiave primaria A e chiave addizionale C, date le dipendenze funzionali $A \rightarrow BC$ e $C \rightarrow A$, la relazione è

- (1) In forma normale di Boyce-Codd (BCNF) e in terza forma normale (3NF)
- (2) In 3NF ma non in BCNF
- (3) Né in 3NF, né in BCNF
- (4) In BCNF e ma non in 3NF



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 30 Gennaio 2020

Regole dell'esame:

- Il tempo massimo per l'esame è di tre ore, dalle 9.30 alle 12.30.
- Non è possibile utilizzare né appunti, né il libro, né altro materiale.
- **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.
- Occorre avere acquisito almeno metà dei punti tra Esercizio 1 e Esercizio 2 (cioè almeno 7.5 punti).
- Occorre avere acquisito almeno metà dei punti dell'Esercizio 3 (cioè almeno 3.5 punti).
- **Coloro che siano sorpresi (1) a copiare, (2) a far copiare, (3) ad usare il cellulare o (4) a consultare qualsiasi materiale:**
 - **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 ad una ditta che estrae marmo per abitazioni da cave, evitando di introdurre entità non necessarie. Disegnare il diagramma ER nel riquadro della pagina che segue.

Si vuole progettare lo schema ER di una base di dati che supporti la gestione di riviste scientifiche e degli articoli pubblicati in esse.

Una rivista scientifica è caratterizzata da un nome univoco, dall'anno di fondazione e dalla tematica trattata (e.g., Computer Science, Architettura, ecc.).

Ciascuna rivista prevede delle edizioni mensili, identificate da un codice univoco e caratterizzate da mese, anno di pubblicazione e prezzo. Ogni edizione mensile è strutturata in pagine (almeno una pagina è sempre presente); una pagina è identificata da una numerazione univoca all'interno dell'edizione.

Ogni articolo è pubblicato in un'unica edizione ed è identificato da un codice univoco all'interno di tale edizione. Di un articolo si conoscono titolo e testo. Ogni articolo occupa un certo numero di pagine (almeno una).

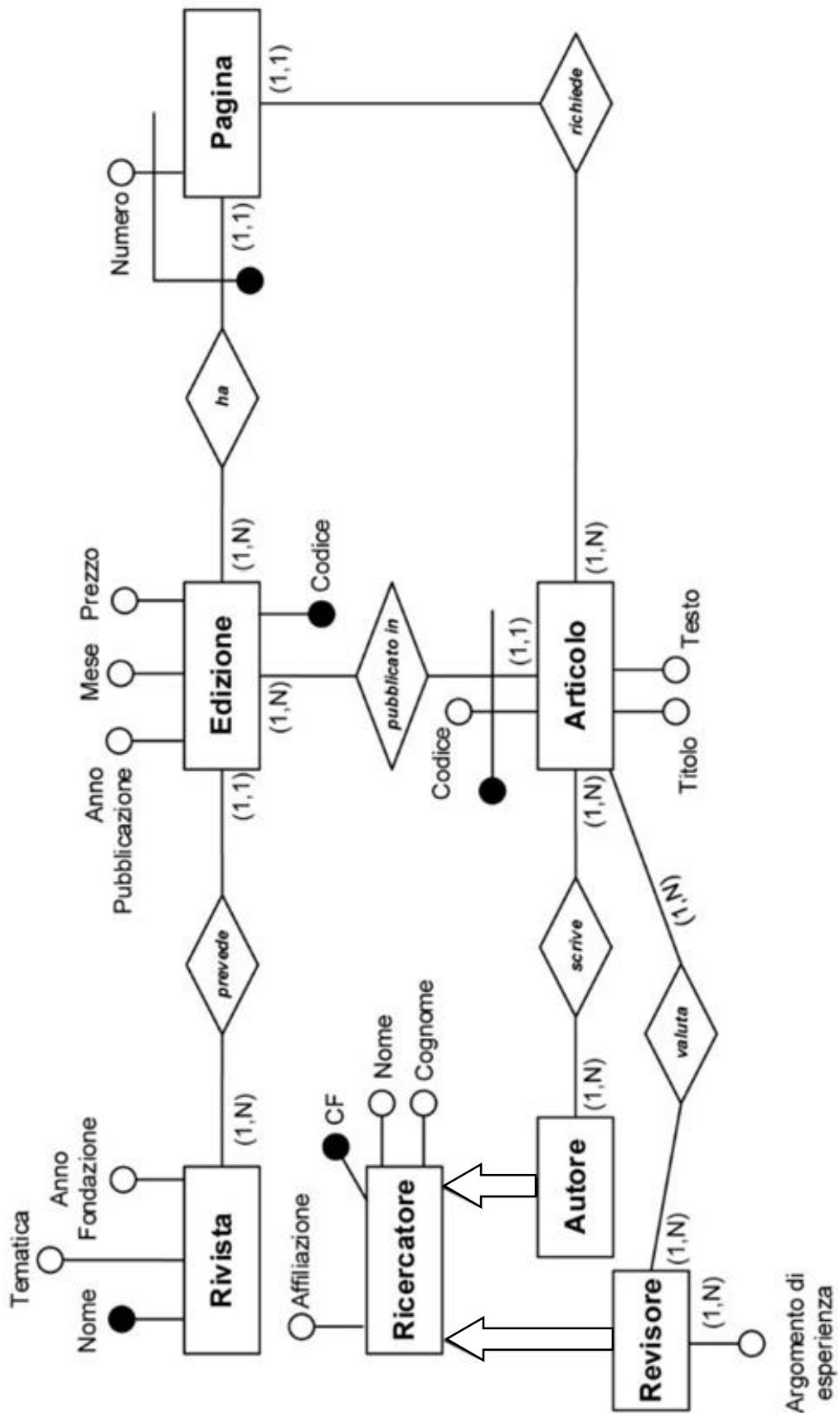
Nella pubblicazione di articoli in riviste scientifiche, sono coinvolti ricercatori. Di ogni ricercatore, si conoscono nome, cognome, codice fiscale, e l'università o centro di ricerca presso la quale afferiscono.

In particolare, alcuni ricercatori sono anche revisori in quanto valutano la qualità degli articoli. Si vuole memorizzare, per ogni articolo, quali revisori (almeno uno) lo hanno revisionato. I ricercatori che sono anche revisori sono associati con un certo numero di argomenti di esperienza.

I ricercatori sono anche autori nel senso che possono pubblicare articoli: per ogni articolo, si vogliono conoscere quali ricercatori ne sono stati autore.

Si noti che un revisore può essere anche un autore, ed un autore può essere anche revisore. Tuttavia, è possibile che un revisore non sia autore, o viceversa.

Soluzione Esercizio 1



Nome: _____ Num. Matric.: _____

Esercizio 2: Progettazione Logica (7 punti)

A partire dallo ER concettuale dell'Esercizio 1, produrre uno schema relazionale del database nel riquadro sottostante, minimizzando i valori nulli. Indicare la chiave primaria, i vincoli di chiave esterne e gli attributi che ammettono valori nulli. Illustrare come ristrutturare il diagramma ER per essere direttamente traducibile in uno schema relazionale.¹

Soluzione Esercizio 2

Nella ristrutturazione:

- l'entità *Ricercatore* viene rimossa e i suoi attributi duplicati per le due entità figlie.
- l'attributo multivalore *Argomento di esperienza* viene tradotto come segue dove l'entità ARGOMENTO contiene un solo attributo *Argomento*, che funge da identificato:



RIVISTA (Nome, Tematica, Anno)

EDIZIONE(Codice, Anno, Mese, Prezzo, NomeRivista)

- Chiave Esterna: EDIZIONE.NomeRivista → RIVISTA.Nome

ARTICOLO(Codice, Edizione, Titolo, Testo)

- Chiave Esterna: ARTICOLO.Edizione → EDIZIONE.Codice

PAGINA(Numero, Edizione, CodiceArt)

- Chiave Esterna: PAGINA.Edizione → EDIZIONE.Codice
- Chiave Esterna: PAGINA.(Edizione,CodiceArt) →

ARTICOLO.(Edizione,Codice)

AUTORE(CF, Affiliazione, Nome, Cognome)

REVISORE(CF, Affiliazione, Nome, Cognome)

SCRIVE(Autore, Codice, Edizione)

- Chiave Esterna: SCRIVE.Autore → Autore.CF
- Chiave Esterna: SCRIVE.(Codice, Edizione) →

ARTICOLO.(Codice, Edizione)

VALUTA(Revisore, Codice, Edizione)

- Chiave Esterna: VALUTA.Revisore → REVISORE.CF
- Chiave Esterna: VALUTA.(Codice, Edizione) →

ARTICOLO.(Codice, Edizione)

ARGOMENTO(Argomento)

ESPERIENZA(Revisore, Argomento)

- Chiave Esterna: ESPERIENZA.Argomento → ARGOMENTO.Argomento
- Chiave Esterna: ESPERIENZA.Revisore → REVISORE.CF

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 per la registrazione degli impiegati che lavorano in aziende:

- IMPIEGATO(CF, Nome, Cognome, Dipartimento, Stipendio)
- LAVORA(CF, PIVA, DataInizio, DataFine)
- AZIENDA(PIVA, Citta, Nazione, Continente)

dove DataFine può essere NULL se l'impiegato lavora ancora in una azienda

- A. Nel riquadro, scrivere una Query in Algebra Relazione che restituisce il codice fiscale degli impiegati che hanno lavorato in almeno due nazioni differenti. (2 punti).²

LA1 = AZIENDA \bowtie LAVORA

LA2 = AZIENDA \bowtie LAVORA

$\Pi_{LA1.CF}$

$(LA1 \bowtie LA1.CF = LA2.CF \text{ AND } LA1.Nazione \neq LA2.Nazione) \text{ LA2}$

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

Nome: _____ Num. Matric.: _____

B. Nel riquadro, scrivere una query SQL che restituiscia il codice fiscale delle persone che attualmente lavorano in esattamente un'azienda, cioè escludendo anche coloro che attualmente non lavorano. Ogni CF può apparire al più una volta (2.5 punti).

```
SELECT DISTINCT CF
FROM LAVORA
WHERE DATAFINE IS NOT NULL
EXCEPT
SELECT L1.CF
FROM LAVORA L1, LAVORA L2
WHERE
    L1.MATR=L2.MATR AND L1.PIVA<>L2.PIVA
    AND L1.DATAFINE IS NULL
    AND L2.DATAFINE IS NULL
```

C. Nel riquadro, scrivere una query SQL che restituisce il cognome degli impiegati che lavorano o hanno lavorato in tutte le aziende. (2.5 punti)

Una soluzione giusta è:

```
CREATE VIEW CF-LAVORA-OVUNQUE (CF) AS
SELECT CF
FROM (SELECT DISTINCT CF, PIVA FROM LAVORA)
GROUP BY CF
HAVING COUNT(*) =
    (SELECT COUNT(*) FROM AZIENDA )
```

```
SELECT COGNOME
FROM IMPIEGATO I, CF-LAVORA-OVUNQUE C
WHERE I.CF=C.CF
```

Tuttavia per l'esame va bene se in questa Query si assume che ogni persona non possa lavorare più di una volta presso una stessa azienda:

```
CREATE VIEW CF-LAVORA-OVUNQUE (CF) AS
SELECT CF FROM LAVORA GROUP BY CF
HAVING COUNT(*) = (SELECT COUNT(*) FROM AZIENDA )
```

Esercizio 4: Normalizzazione (6 punti)

Data la relazione R(A,B,C,D) con dipendenze funzionali
 $\{ C \rightarrow D, \quad C \rightarrow A, \quad B \rightarrow C \}$.

1. Mostrare tutte le chiavi di R e motivare perché ognuna è chiave.
2. Dire quale dipendenze violano la forma normale di Boyce Codd (BCNF), spiegandone la ragione.
3. Decomporre in BCNF

1. $B^+ = \{B, C, A, D\}$ e $C^+ = \{C, A, D\}$. Quindi, B è chiave perché la sua chiusura contiene tutti gli attributi della relazione
2. $C \rightarrow D$ e $C \rightarrow A$ violano BCNF perché C non è una superchiave della relazione
3. È possibile decomporre come segue: $R_1(\underline{C}, D)$, $R_2(\underline{C}, A)$ e $R_3(\underline{B}, C)$. Eventualmente – non necessario per l'esame – è possibile rimettere insieme R_1 e R_2 , ottenendo la decomposizione: $R_4(\underline{C}, A, D)$, $R_3(\underline{B}, C)$.

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 R WHERE C='Valore'` sulla relazione $R(\underline{A}, B, C)$. Quale dei seguenti indici in genere assicura le migliori performance in termini di velocità dell'esecuzione della query?

- (1) Indice Hash su C;
- (2) Indice B-Tree su C;³
- (3) Indice Hash su A, B, C;
- (4) Indice B-Tree su A, B, C.

Domanda 2 (1.5 Punti)

Date una qualsiasi istanza delle 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 ci sono 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 $R \bowtie_{B=F} S$?

- (1) $0 \leq |R \bowtie_{B=F} S| \leq |R|$
- (2) $0 \leq |R \bowtie_{B=F} S| \leq |R| * |S|$
- (3) $|R \bowtie_{B=F} S| = |R| * |S|$
- (4) $|R| \leq |R \bowtie_{B=F} S| \leq |R| * |S|$

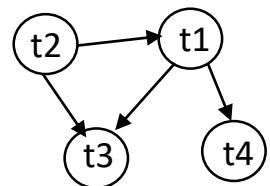
³ Nota sull'esercizio: la risposta "B-Tree su C" non è del tutto sbagliata perché comunque quell'indice velocizza la query, anche se non tanto quanto quello Hash.

Domanda 3 (2 Punti)

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

R₂(A) R₂(B) W₁(A) W₂(D) R₃(C) R₁(C) W₃(B) R₄(A) 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-serializzabile⁴
- (4) S **non è né** conflict-serializzabile **né** view-serializzabile

⁴ Nota sulla soluzione della Domanda 3: non è necessario verificare la view-serializzabilità perché la teoria dice che, se lo schedule è conflict-serializzabile, allora è anche view-serializzabile.



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 24 Febbraio 2020

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 ad una catena di **pista di pattinaggio**, evitando di introdurre entità non necessarie. Disegnare il diagramma ER nel riquadro della pagina che segue.*

Il sistema deve memorizzare le informazioni di ogni pista della catena. Di ogni pista, è di interesse sapere (1) la città in cui si trova, (2) l'anno di apertura e (3) il fatturato. Ogni città non può avere più di una pista. Inoltre, ogni pista è gestita da una persona, però una pista potrebbe temporaneamente non avere una persona che la gestisce. Si noti che una persona non può prendere in gestione più di una pista della catena.

Il sistema memorizza un'anagrafe delle persone: di ogni persona, è di interesse sapere il codice fiscale (identificativo), la data di nascita e la città di nascita.

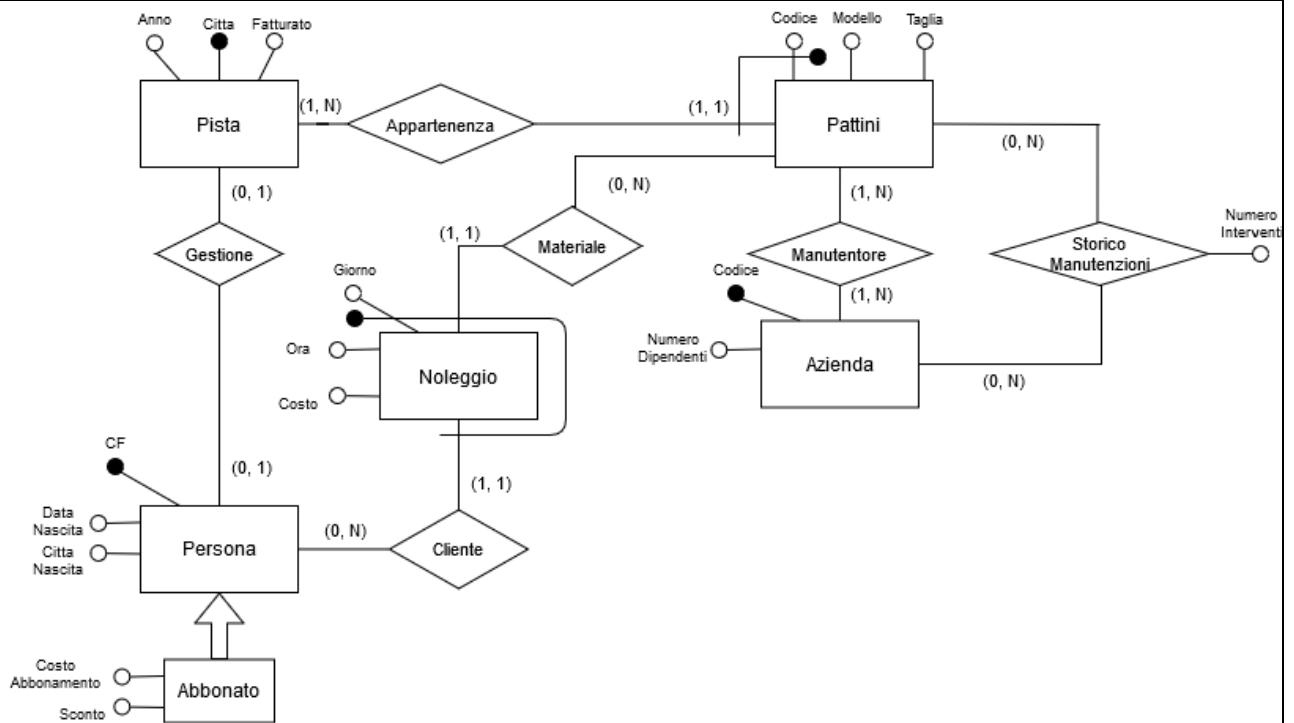
Quando le persone desiderano essere clienti di una pista, noleggiano un paio di pattini. Ogni paio di pattini è associato ad esattamente una pista.

Di ogni paio di pattini interessa il codice (unico all'interno della pista), il modello, la taglia. Inoltre, per ogni paio di pattini, è di interesse sapere quali aziende sono in grado di farne manutenzione. Si vuole anche memorizzare lo storico delle manutenzioni: per ogni paio di pattini, è di interesse sapere quante volte (intero) il paio è andato in manutenzione in ogni azienda in grado di manutenere. Di ogni azienda di riparazione pattini interessa il codice identificativo e il numero di dipendenti.

Come detto, ogni paio di pattini è noleggiato da persone per usarle nella pista associata. In particolare, si vuole sapere per ogni noleggio: (1) quale paio è stato preso, (2) il giorno e l'ora del noleggio e (3) il costo complessivo del noleggio. Si noti che ogni paio di pattini può essere noleggiato più volte dalla stessa persona, ma solo in giorni diversi.

Alcuni persone sono clienti abbonati. Per tali clienti, è di interesse sapere il costo dell'abbonamento annuale (unico per cliente) e quale sconto (in percentuale) ottengono nei noleggi in virtù del loro abbonamento. Si noti che, se una persona è abbonata, non può prendere in gestione una pista (occorre modellare questo vincolo esplicitamente nello ER). Viceversa, una persona che non è abbonata può noleggiare pattini ed essere gestore di una pista. Non è necessario modellare il vincolo che non è possibile noleggiare pattini relativi all'eventuale pista in gestione.

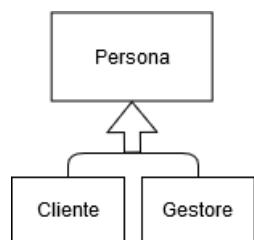
Soluzione Esercizio 1



Di seguito sono alcuni punti della soluzione che meritano un approfondimento:

- L'identificatore dell'entità **Pista** è la città, poiché non è possibile avere due piste nella stessa città. L'aggiunta di un codice come identificatore è errato perché non è previsto dalle specifiche, né necessario. Inoltre, tale codice come identificatore non permette di rappresentare il vincolo che non è possibile avere due istanze dell'entità **Pista** con lo stesso valore per l'attributo Città.
- Il codice non è sufficiente come identificatore dell'entità **Pattini** perché è unico per i pattini di una pista, ma non unico per tutti i pattini della catena.
- Si noti l'identificatore dell'entità **Noleggio**. L'identificatore modella che non è possibile che la stessa persona noleggi gli stessi pattini nello stesso giorno. Si osservi che l'ora non è parte dell'identificatore, altrimenti sarebbe possibile noleggiare i pattini ad orari diversi dello stesso giorno.
- È errato realizzare *Noleggio* come relazione tra *Persona* e *Pattini* invece di entità come proposto sopra.** Se *Noleggio* fosse una relazione, non sarebbe possibile noleggiare gli stessi pattini più di una volta dalla parte della stessa persona: il testo indica che tale eventualità è possibile ma in giorni diversi.

L'ultima osservazione riguarda la possibilità qui a fianco (con **Persona** che generalizza le entità **Cliente** e **Gestore**). Tale soluzione è errata. La generalizzazione è esclusiva: vorrebbe dire che una persona può essere solamente cliente oppure gestore. Si noti che il fatto che la generalizzazione sia totale (freccia piena) o parziale (freccia vuota) non cambia nulla.



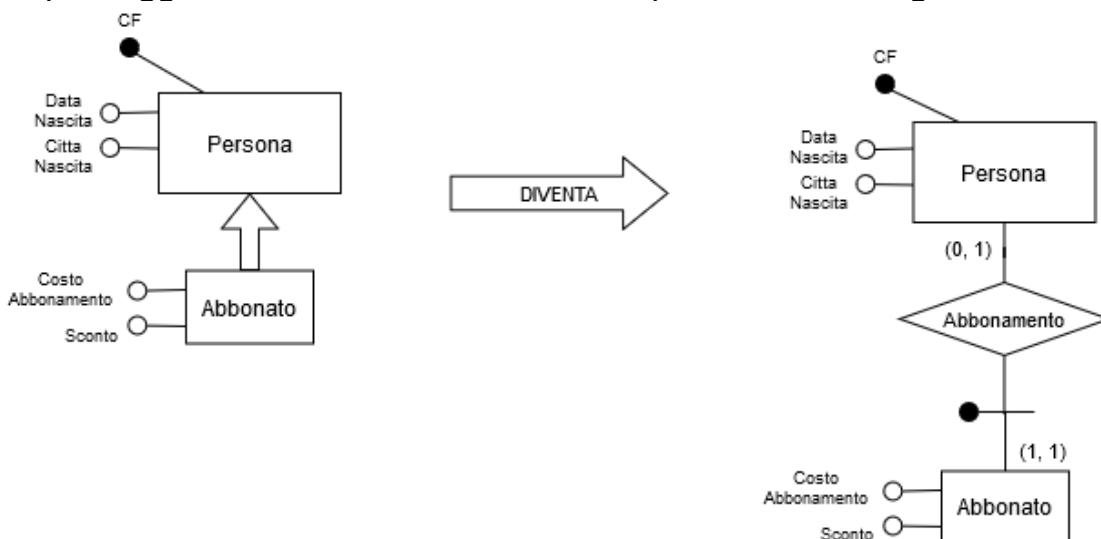
Nome: _____ Num. Matric.: _____

Esercizio 2: Progettazione Logica (7 punti)

A partire dallo ER concettuale dell'Esercizio 1, produrre uno schema relazionale del database nel riquadro sottostante, minimizzando i valori nulli. Indicare la chiave primaria, i vincoli di chiave esterne e gli attributi che ammettono valori nulli. Illustrare come ristrutturare il diagramma ER per essere direttamente traducibile in uno schema relazionale.¹

Soluzione Esercizio 2

Il primo passaggio è la ristrutturazione dello ER per rimuovere la generalizzazione:



Alla luce di quanto sopra, lo schema logico è il seguente:

PERSONA(CF, DataNascita, CittaNascita)

PISTA(Città, Fatturato, Anno, Gestore)

- Chiave Esterna: PISTA.Gestore → Persona.CF
- Chiave Aggiuntiva: Gestore

PATTINI(Codice, CittaPista, Modello, Taglia)

- Chiave Esterna: PATTINI.CittaPista → PISTA.Città

NOLEGGIO(CodicePattini, CittaPista, Persona, Giorno, Ora, Costo)

- Chiave Esterna: NOLEGGIO.(CodicePattini, CittaPista) → PATTINI.(Codice, CittaPista)
- Chiave Esterna: NOLEGGIO.Persona → Persona.CF

AZIENDA(Codice, NumeroDipendenti)

MANUTENTORE(CodicePattini, CittaPista, CodiceAzienda)

- Chiave Esterna: MANUTENTORE.(CodicePattini, CittaPista) → PATTINI.(Codice, CittaPista)
- Chiave Esterna: MANUTENTORE.CodiceAzienda → AZIENDA.Codice

STORICO(CodicePattini, CittaPista, CodiceAzienda, NumeroInterventi)

- Chiave Esterna: STORICO.(CodicePattini, CittaPista) → PATTINI.(Codice, CittaPista)
- Chiave Esterna: STORICO.CodiceAzienda → AZIENDA.Codice

ABBONATO(CF, Costo, Sconto)

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 per la registrazione di libri e dei loro autori, dove i libri sono identificati dal nome:

- AUTORE(CF, Nome, Eta)
- PUBBLICA(CF, Libro)

A. Nel riquadro, scrivere una Query in Algebra Relazione che restituisce il codice fiscale e il nome degli autori che hanno pubblicato almeno due libri. (2 punti).²

P1 = PUBBLICA

P2 = PUBBLICA

$$\begin{aligned} \Pi_{\text{AUTORE.CF}, \text{AUTORE.NOME}} \\ (\text{AUTORE} \bowtie_{\text{AUTORE.CF} = \text{P1.CF}} \\ (\text{P1} \bowtie_{\text{P1.CF} = \text{P2.CF} \text{ AND } \text{P1.Libro} <> \text{P2.Libro}} \text{P2})) \end{aligned}$$

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

Nome: _____ Num. Matric.: _____

B. Nel riquadro, scrivere una query SQL per calcolare la media dell'età degli autori che hanno scritto almeno un libro con altri (cioè che hanno scritto almeno un libro che hanno più di un autore). (2.5 punti).

```
SELECT AVG(ETA),
FROM AUTORE
WHERE CF IN
  (SELECT P1.CF
   FROM PUBBLICA P1, PUBBLICA P2
   WHERE P1.LIBRO=P2.LIBRO
   AND P1.CF <> P2.CF)
```

C. Nel riquadro, scrivere una query SQL che restituisce il libro o i libri con il numero più alto di autori. (2.5 punti)

```
CREATE VIEW AUTORI-PER-LIBRO
(LIBRO,NUM_AUTORI) AS
```

```
  SELECT LIBRO, COUNT(*)
  FROM PUBBLICA
 GROUP BY LIBRO
```

```
SELECT LIBRO
FROM AUTORI-PER-LIBRO
WHERE NUM_AUTORI=
  (SELECT MAX(NUM_AUTORI)
   FROM AUTORI-PER-LIBRO)
```

Esercizio 4: Transazioni (5 punti)

Data la transazione $S = r_1(x)w_2(x)r_3(x)w_1(u)w_3(v)r_3(y)r_2(y)w_3(u)w_4(t)w_3(t)$, sapendo che S è conflict-serializzabile:

1. Mostrare/Spiegare se S è (o non è) view-serializzabile;
2. se S view-serializzabile, mostrare uno schedule seriale T che è view-equivalente a S , mostrando perché S and T sono view-equivalenti.

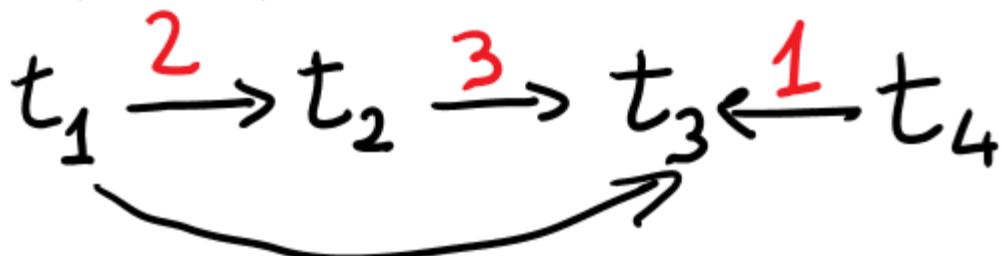
Parte 1

Siccome S è conflict-serializzabile, è anche view-serializzabile.

Parte 2

Se S è conflict-serializzabile, è possibile costruire un grafo dei conflitti di S che è aciclico. Il percorso nel grafo che tocca tutte le transazioni, fornisce un ordine delle transazioni. Scrivendo le operazioni nell'ordine delle transazioni e rispettando l'ordine all'interno delle transazioni, si ottiene uno schedule seriale che è conflict-equivalente e quindi anche view-equivalente.

Il seguente è il grafo dei conflitti:



Quindi, è possibile serializzare le transizioni nel seguente ordine: t_1, t_2, t_4, t_3 . Questo corrisponde allo schedule seriale $T = r_1(x) w_1(u) w_2(x) r_2(z) w_4(t) r_3(x) w_3(v) r_3(z) w_3(u) w_3(t)$, che è conflict-equivalente e quindi anche view-equivalente.

Si poteva ovviamente anche lavorare direttamente sulla view-equivalenza, guardando alla relazione leggi-da e alle scritture finali. In S , $r_3(x)$ legge da $w_2(x)$, mentre $r_1(x)$, $r_2(x)$ e $r_3(z)$ non leggono da scritture di transazioni in S . Le scritture finali sono $w_2(x)$, $w_3(t)$, $w_3(u)$ e $w_3(v)$. È facile verificare che lo schedule T ha una sola leggi-da come S , cioè $r_3(x)$ legge da $w_2(x)$, e le stesse scritture finali.

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 relazione $R(A, B, C)$ con dipendenze funzionali $A \rightarrow BC$ e $B \rightarrow A$. Una delle seguenti affermazioni è vera. Quale?

- (1) AB è superchiave **ma non è chiave**; (Soluzione Corretta)
- (2) AB è **sia** chiave **che** superchiave;
- (3) AB **non è né** chiave **né** superchiave;
- (4) AB è chiave **ma non è** superchiave.

Domanda 2 (1 Punti)

Data un'istanza della relazione $R(A, B, C)$ dove l'unica chiave di R è {A, C}, (2) nessun attributo ammette valori nulli . Indicato con $|X|$ il numero di tuple di una relazione X, quale è vera tra le seguenti affermazioni relative al numero di tuple in $\pi_A(R)$?

- (1) È possibile sia $|\pi_A(R)| < |R|$ che $|\pi_A(R)| = |R|$. (Sol. Corretta)
- (2) È sempre il caso che $|\pi_A(R)| = |R|$.
- (3) Non è possibile che $|\pi_A(R)| = |R|$.
- (4) È sempre il caso che $|\pi_A(R)| = |\pi_{AC}(R)|$.

Domanda 3 (2.5 Punti)

Sia data una relazione $S(ABCD)$ con il set di dipendenze funzionali $\{C \rightarrow A, C \rightarrow D, D \rightarrow B\}$. Dopo aver trovato la chiave (che è una), dire quale affermazione (una) tra le seguenti è vera.

- (1) S è in formale normale di Boyce and Codd (BCNF) ma **non è** in terza forma normale (3NF).
- (2) S è **sia** in BCNF **che** in 3NF.
- (3) S **non è** in BCNF **ma è** in 3NF.
- (4) S **non è né** in BCNF **né** in 3NF. (Soluzione corretta con C unica chiave)



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Seconda Prova dell'esame on-line di Maggio (data 16 Maggio)

Progettazione Concettuale (8 punti) e Logica (7 punti)

Si richiede di produrre i seguenti artefatti:

- **Schema Concettuale Entità-Relazione**, da inviare per email da account UniPd)
- **Schema Logico**, da inserire nel riquadro di cui sotto. Indicare la chiave primaria, i vincoli di chiavi esterne e gli attributi che ammettono valori nulli.

per un'applicazione relativa ad una catena di piste di pattinaggio. Nella progettazione concettuale e logica, occorre evitare di introdurre entità e tavole non necessarie ed occorre minimizzare i valori nulli.

Illustrare come ristrutturare il diagramma ER per essere direttamente traducibile in uno schema relazionale. Il diagramma ER ristrutturato deve anche essere inviato per email da account UniPd.

Il sistema deve memorizzare le informazioni di ogni pista della catena. Di ogni pista, è di interesse sapere (1) la città in cui si trova, (2) l'anno di apertura e (3) il fatturato. Ogni città non può avere più di una pista. Inoltre, ogni pista è gestita da una persona, però una pista potrebbe temporaneamente non avere una persona che la gestisce. Si noti che una persona non può prendere in gestione più di una pista della catena.

Il sistema memorizza un'anagrafe delle persone: di ogni persona, è di interesse sapere il codice fiscale (identificativo), la data di nascita e la città di nascita.

Quando le persone desiderano essere clienti di una pista, noleggiano un paio di pattini. Ogni paio di pattini è associato ad esattamente una pista.

Di ogni paio di pattini interessa il codice (unico all'interno della pista), il modello, la taglia. Inoltre, per ogni paio di pattini, è di interesse sapere quali aziende sono in grado di farne manutenzione. Si vuole anche memorizzare lo storico delle

manutenzioni: per ogni paio di pattini, è di interesse sapere quante volte (intero) il paio è andato in manutenzione in ogni azienda in grado di manutenere. Di ogni azienda di riparazione pattini interessa il codice identificativo e il numero di dipendenti.

Come detto, ogni paio di pattini è noleggiato da persone per usarle nella pista associata. In particolare, si vuole sapere per ogni noleggio: (1) quale paio è stato preso, (2) il giorno e l'ora del noleggio e (3) il costo complessivo del noleggio. Si noti che ogni paio di pattini può essere noleggiato più volte dalla stessa persona, ma solo in giorni diversi.

Alcuni persone sono clienti abbonati. Per tali clienti, è di interesse sapere il costo dell'abbonamento annuale (unico per cliente) e quale sconto (in percentuale) ottengono nei noleggi in virtù del loro abbonamento.

(8 punti per la modellazione concettuale + 7 punti per lo schema logica)

Notazione per lo schema logico:

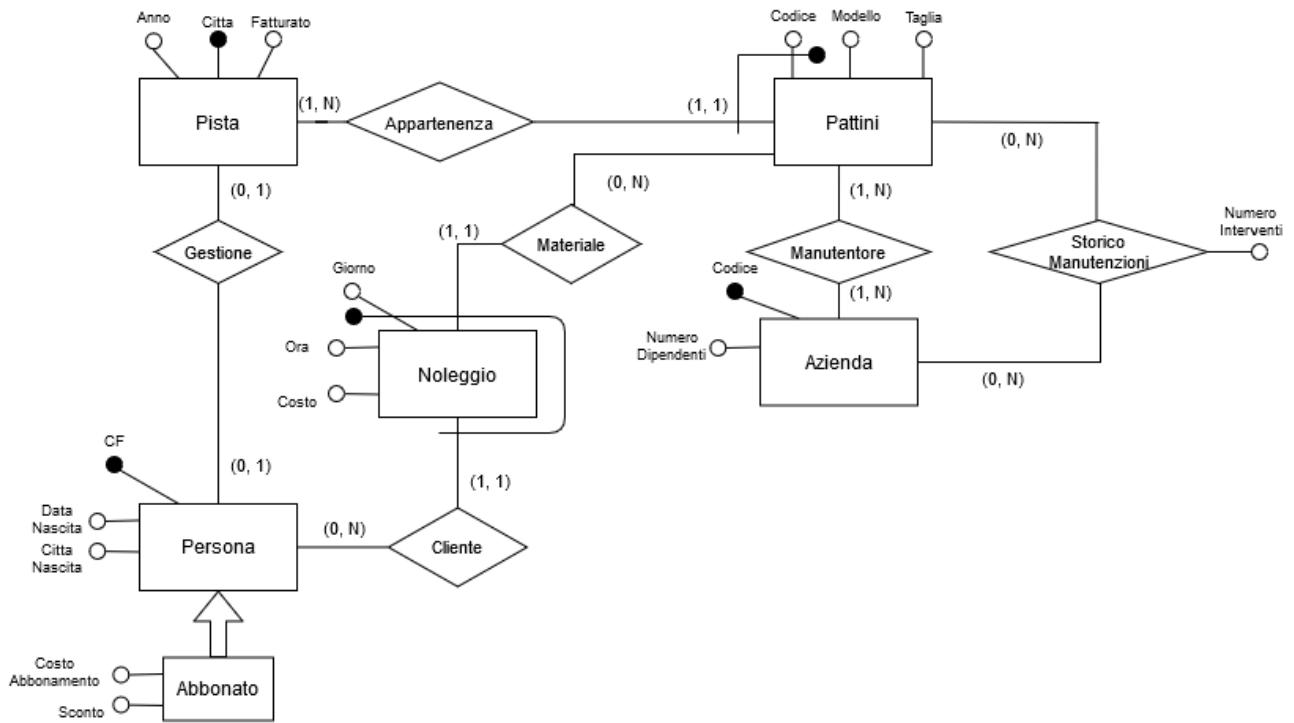
Usare il sottolineato per indicare i vincoli di chiave primaria, e il **bold** per indicare gli attributi che ammettono valori nulli.

Indicare con X.A-->Y.B per indicare che l'attributo A della tabella X è chiave esterna all'attributo B della tabella Y.

Indicare con X.(A,B) --> Y.(C,D) per indicare che gli attributi (A,B) della tabella X sono chiavi esterne all'attributo (C,D) della tabella Y.

Soluzione Progettazione Concettuale

Una soluzione è data dal seguente schema ER:

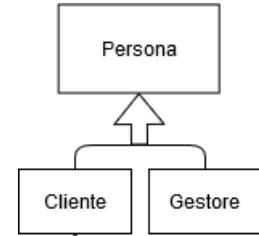


Di seguito sono alcuni punti della soluzione che meritano un approfondimento:

- L'identificatore dell'entità **Pista** è la città, poiché non è possibile avere due piste nella stessa città. L'aggiunta di un codice come identificatore è errato perché non è previsto dalle specifiche, né necessario. Inoltre, tale codice come identificatore non permettere di rappresentare il vincolo che non è possibile avere due istanze dell'entità **Pista** con lo stesso valore per l'attributo Città.
- Il codice non è sufficiente come identificatore dell'entità **Pattini** perché è unico per i pattini di una pista, ma non unico per tutti i pattini della catena.
- Si noti l'identificatore dell'entità **Noleggio**. L'identificatore modella che non è possibile che la stessa persona noleggi gli stessi pattini nello stesso giorno. Si osservi che l'ora non è parte dell'identificatore, altrimenti sarebbe possibile noleggiare i pattini ad orari diversi dello stesso giorno.
- **È errato realizzare *Noleggio* come relazione tra *Persona* e *Pattini* invece di entità come proposto sopra.** Se **Noleggio** fosse una relazione, non sarebbe possibile noleggiare gli stessi pattini più di una volta dalla parte della stessa persona: il testo indica che tale eventualità è possibile ma in giorni diversi.
- Si è considerato accettabile omettere la relazione **Manutentore**, assumendo che numero di interventi è 0 quando una azienda è registrata per la manutenzione, ma non ha effettuato ancora interventi. In tal caso, comunque,

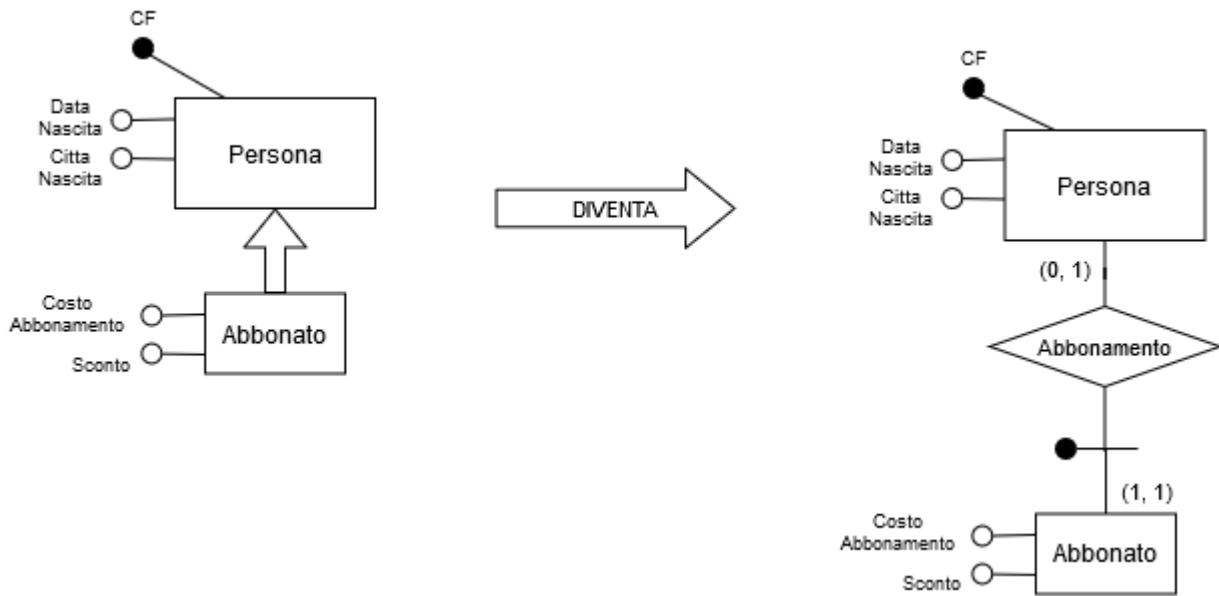
la cardinalità della relazione *Storico Manutentori* deve essere (1,N) dalla parte dell'azienda: ogni azienda è inserita se e solo se c'è almeno un paio di pattini che può essere manutenuta da essa.

L'ultima osservazione riguarda la possibilità qui a fianco (con *Persona* che generalizza le entità *Cliente* e *Gestore*). Tale soluzione è errata a meno che viene esplicitamente detto che la generalizzazione persona-cliente-gestore è sovrapposta. Se non viene detto che è *sovrapposta* viene considerata il default, che è *esclusiva*: vorrebbe dire che una persona può essere solamente cliente oppure gestore. Si noti che il fatto che la generalizzazione sia totale (freccia piena) o parziale (freccia vuota) non cambia nulla.



Soluzione “Progettazione Logica”

Il primo passaggio è la ristrutturazione dello ER per rimuovere la generalizzazione:



Si è deciso in questo caso di sostituire la generalizzazione con la relazione *Abbonamento*. La soluzione alternativa sarebbe stata quella di inglobare l'entità *Abbonato* all'interno di *Persona*. Tuttavia, per tutti coloro che non sono abbonati, gli attributi *Costo Abbonamento* e *Sconto* sarebbero dovuti essere nulli. Come indicato nel testo, si vuole minimizzare il numero di valori nulli.

Alla luce di quanto sopra, lo schema logico è il seguente:

PERSONA(CF, DataNascita, CittaNascita)

PISTA(Citta, Fatturato, Anno, **Gestore**)

- Chiave Esterna: PISTA.Gestore → Persona.CF
- Chiave Aggiuntiva: Gestore¹

PATTINI(Codice, CittaPista, Modello, Taglia)

- Chiave Esterna: PATTINI.CittaPista → PISTA.Citta

NOLEGGIO(CodicePattini, CittaPista, Persona, Giorno, Ora, Costo)

- Chiave Esterna: NOLEGGIO.(CodicePattini, CittaPista) → PATTINI.(Codice, CittaPista)
- Chiave Esterna: NOLEGGIO.Persona → Persona.CF

AZIENDA(Codice, NumeroDipendenti)

MANUTENTORE(CodicePattini, CittaPista, CodiceAzienda)

- Chiave Esterna: MANUTENTORE.(CodicePattini, CittaPista) → PATTINI.(Codice, CittaPista)
- Chiave Esterna: MANUTENTORE.CodiceAzienda → AZIENDA.Codice

STORICO(CodicePattini, CittaPista, CodiceAzienda, Numerointerventi)

- Chiave Esterna:² STORICO.(CodicePattini, CittaPista) → PATTINI.(Codice, CittaPista)
- Chiave Esterna:² STORICO.CodiceAzienda → AZIENDA.Codice

ABBONATO(CF, Costo, Sconto)

- Chiave Esterna: ABBONATO.CF → PERSONA.CF

Ad essere rigorosi nel minimizzare i valori nulli, occorre:

- avere la relazione *PISTA(Citta, Fatturato, Anno)*,
- introdurre una relazione *GESTIONE(CittaPista, Gestore)* dove
GESTIONE.CittaPista → PISTA.Citta e GESTIONE.Gestore → PERSONA.CF

Questo garantisce assenza di NULL: una tupla (citPist, CF) esiste in GESTIONE se la pista nella città citPist è gestita dalla persona con codice fiscale CF. Similmente alla prima soluzione occorre aggiungere che *Gestore* è anche chiave (non primaria).

¹ Questa chiave aggiuntiva è necessaria per assicurare che ci sia al più un solo gestore. Tuttavia non era richiesta aggiungere altra chiave oltre a quella primaria. Quindi, se non è stata inserita non è considerato errore.

² Volendo arrivare ad una soluzione ancora migliore, si può osservare che le tuple (pattini, pista, azienda) in storico sono un sottoinsieme di quelli in manutentore:

STORICO.(CodicePattini, CittaPista, CodiceAzienda) → MANUTENTORE.(CodicePattini, CittaPista, CodiceAzienda)

Assegnamento del Punteggio

Parte del Dominio di Modellazione	Modellazione Concettuale					Modellazione Logica (tabelle)					
	Punti per la modellazione	Errori Commessi			Punti per la modellazione	Errori Commessi			1	2	3+
		1	2	3+		1	2	3+			
Modellazione delle Piste e del Gestore (Entità Pista e Relazione Gestione)	1 punto	-0.5 punti	-1 punto	-1 punto	0.5 punti	-0.25 punti	-0.5 punti	-0.5 punti			
Modellazione dei Pattini e della loro Appartenenza a Piste (Pattini e Appartenenza)	1.5 punti	-0.5 punti	-1 punto	-1.5 punto	1.5 punti	-0.5 punti	-1 punto	-1.5 punti			
Modellazione dei Noleggi (Noleggio, Materiale, Cliente) (L'errore di considerare "Noleggio" come relazione corrisponde di fatto a molteplici errori con una penalizzazione totali di punti: 1.5)	1.5 punti	-0.75 punti	-1.25 punti	-1.5 punti	1.5 punti	-0.5 punti	-1 punto	-1.5 punti			
Modellazione delle Manutenzioni (Manutentore, Storico e Azienda)	2 punti	-0.75 punti	-1.5 punti	-2 punti	1.25 punti	-0.5 punti	-1 punto	-1.25 punti			
Modellazione della Persona e dell'Abbonato (L'errore sulla generalizzazione nello ER concettuale è grave e corrisponde a una penalizzazione totali di punti: 2)	2 punti	-0.75 punti	-1.5 punti	-2 punti	1.25 punti	-0.5 punti	-1 punto	-1.25 punti			
Conversione allo ER "ristrutturato" <i>(L'errore nella conversione della generalizzazione è grave e corrisponde a una penalizzazione totali di punti: 1)</i>	Non applicabile					1	0.5	0	0		

Esercizio Transazioni (5 punti)

Data lo schedule $S = r1(x)w2(x)r3(x)w1(u)w3(v)r3(y)r2(y)w3(u)w4(t)w3(t)$ sapendo che S è conflict-serializzabile:

1. Mostrare/Spiegare se S è (o non è) view-serializzabile;
2. se S view-serializzabile, mostrare uno schedule seriale T che è view-equivalente a S , mostrando perché S and T sono view-equivalenti.

Soluzione

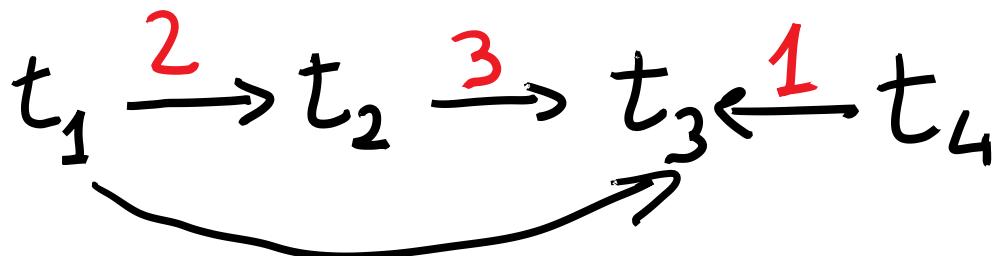
Parte 1

Siccome S è conflict-serializzabile, è anche view-serializzabile.

Parte 2

Se S è conflict-serializzabile, è possibile costruire un grafo dei conflitti di S che è aciclico. Il percorso nel grafo che tocca tutte le transazioni, fornisce un ordine delle transazioni. Scrivendo le operazioni nell'ordine delle transazioni e rispettando l'ordine all'interno delle transazioni, si ottiene uno schedule seriale che è conflict-equivalente e quindi anche view-equivalente.

Il seguente è il grafo dei conflitti:



Quindi, è possibile serializzare le transizioni nel seguente ordine: $t1, t2, t4, t3$.³

Questo corrisponde allo schedule seriale $T = r1(x) w1(u) w2(x) r2(z) w4(t) r3(x) w3(v) r3(z) w3(u) w3(t)$, che è conflict-equivalente e quindi anche view-equivalente.

Si poteva ovviamente anche lavorare direttamente sulla view-equivalenza, guardando alla relazione leggi-da e alle scritture finali. In S , $r3(x)$ legge da $w2(x)$, mentre $r1(x)$, $r2(x)$ e $r3(z)$ non leggono da scritture di transazioni in S . Le scritture finali sono $w2(x)$, $w3(t)$, $w3(u)$ e $w3(v)$. È facile verificare che lo schedule T ha una sola leggi-da come S , cioè $r3(x)$ legge da $w2(x)$, e le stesse scritture finali.

³ In realtà, va bene qualsiasi ordine in cui appaia la sequenza $\langle t1, t2, t3 \rangle$ e $t4$ sia prima di $t3$. Dunque anche: $\langle t4, t1, t2, t3 \rangle$ oppure $\langle t1, t4, t2, t3 \rangle$.

Assegnamento di punteggio

- **1 punto:** la parte 1 (CSR implica VSR), che è banale;

Per la parte 2 da aggiungere alla parte 1:

- **2.5 punti:** Una soluzione corretta con una spiegazione corretta ma vaga (per es. “T ha la stessa relazione leggi-da e le stesse scritture finali” senza indicare quali)
- **5 punti:** Una spiegazione dettagliata (per es. spiegando rispetto al grafo dei conflitti o mostrando la relazione leggi-da ed



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Seconda Prova del Terzo Appello (data 3 Luglio)

Progettazione Concettuale (8 punti) e Logica (7 punti)

Si richiede di produrre i seguenti artefatti:

- **Schema Concettuale Entità-Relazione**, da inviare per email da account UniPd)
- **Schema Logico**, da inserire nel riquadro di cui sotto. Indicare la chiave primaria, i vincoli di chiave esterne e gli attributi che ammettono valori nulli.

per un'applicazione relativa ad un **negozi di fumetti**. Nella progettazione concettuale e logica, occorre evitare di introdurre entità e tabelle non necessarie ed occorre minimizzare i valori nulli.

Illustrare come ristrutturare il diagramma ER per essere direttamente traducibile in uno schema relazionale. **Il diagramma ER ristrutturato deve anche essere inviato per email da account UniPd.**

Il negozio di fumetti vuole costruire un database relativo ai fumetti in vendita nel negozio, tenendo traccia dei fumetti venduti e dei clienti che hanno effettuato acquisti.

Ogni fumetto è identificato dal nome della serie di cui fa parte e dal numero che lo contraddistingue all'interno della serie (ad es. Topolino n.21243). Per tutti i fumetti si vuole inoltre conoscere l'anno di pubblicazione.

Per alcune tipologie di fumetti sono di particolare interesse informazioni addizionali. Per i manga (i fumetti giapponesi) interessa conoscere l'autore, mentre per i fumetti western è interesse conoscere la casa editrice.

Il negozio vuole tener traccia delle varie copie disponibili per ciascun fumetto; ciascuna copia è identificata un numero progressivo e da un prezzo diverse. Copie diverse dello stesso fumetto possono avere prezzi di vendita diverse (per esempio, in funzione del momento di stampa della copia).

Ciascun cliente (di cui interessa conoscere il codice fiscale, il nome, il cognome e l'e-mail) può fare le seguenti cose:

- Acquistare copie di fumetti: in tal caso dell'acquisto si memorizza la data e l'ora dello stesso
- Prenotare fumetti di interesse: un cliente può prenotare un certo numero di copie di un particolare fumetto, memorizzando data della prenotazione. Lo stesso cliente può prenotare lo stesso fumetto più volte, ma solo in date diverse.

Si noti che il negozio può vendere anche copie usate di un fumetto, e in questo caso è di interesse conoscere lo sconto applicato rispetto alla versione venduta a prezzo pieno

(8 punti per la modellazione concettuale + 7 punti per lo schema logica)

Notazione per lo schema logico:

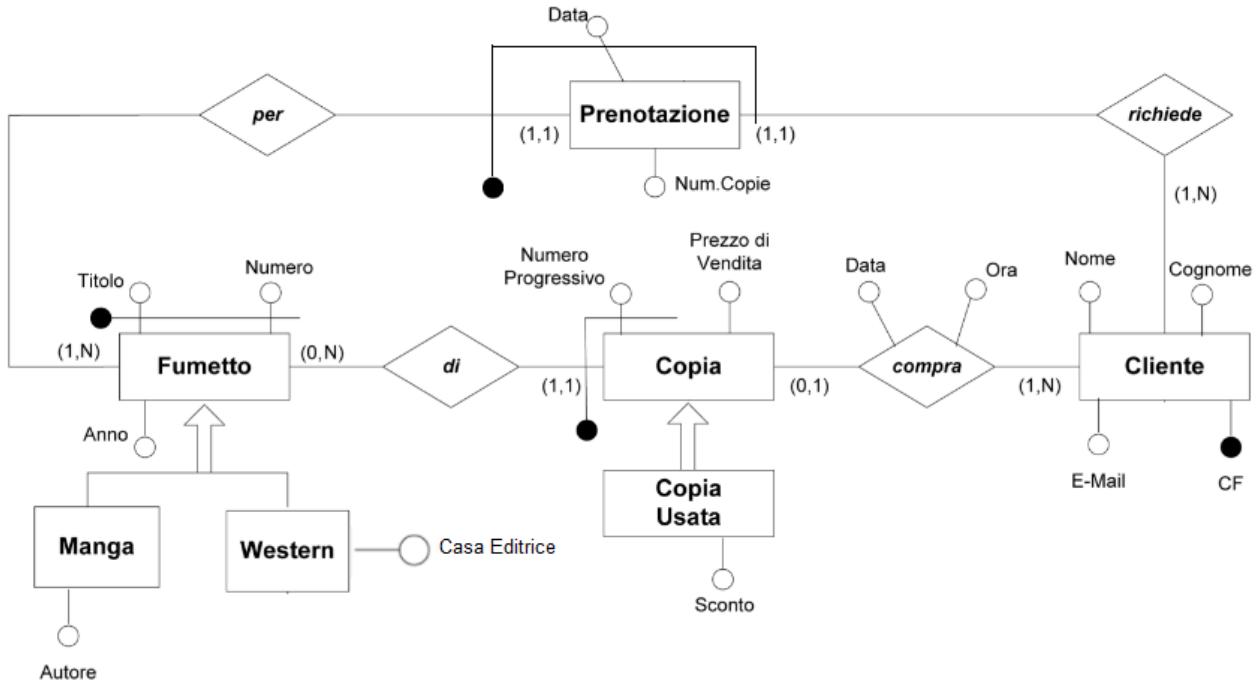
Usare il sottolineato per indicare i vincoli di chiave primaria, e il **bold** per indicare gli attributi che ammettono valori nulli.

Indicare con X.A-->Y.B per indicare che l'attributo A della tabella X è chiave esterna all'attributo B della tabella Y.

Indicare con X.(A,B) --> Y.(C,D) per indicare che gli attributi (A,B) della tabella X sono chiavi esterne all'attributo (C,D) della tabella Y.

Soluzione Progettazione Concettuale

Una soluzione è data dal seguente schema ER:

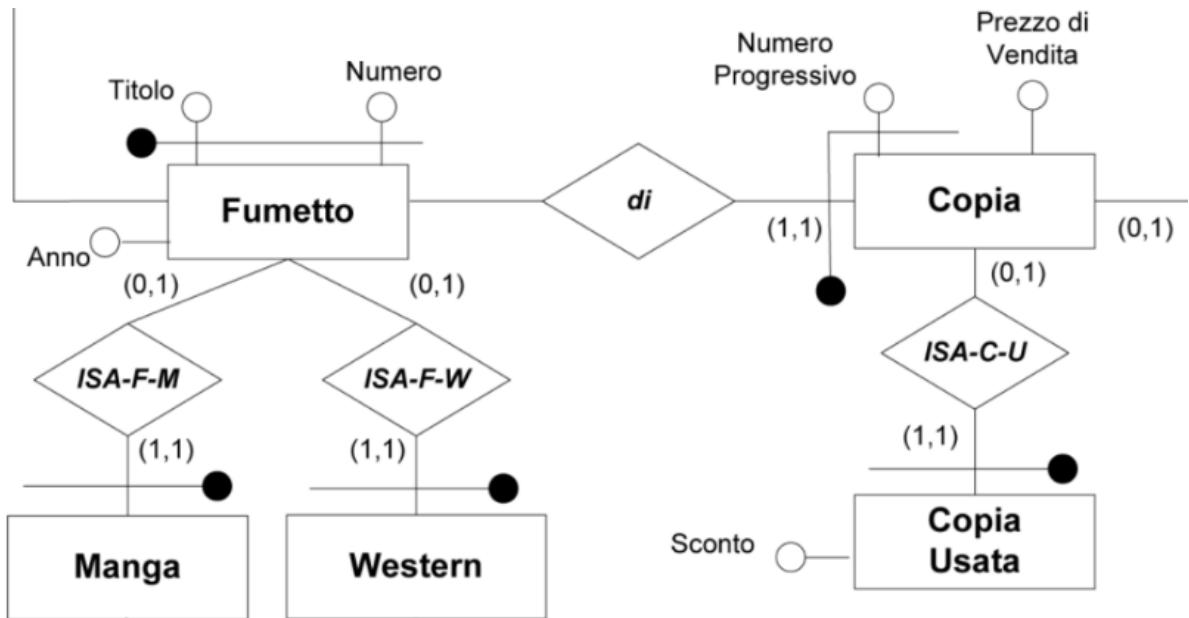


Di seguito sono alcuni punti della soluzione che meritano un approfondimento:

- La generalizzazione tra Fumetto, Manga e Western è parziale (freccia bianca). Ci sono fumetti che non sono né Manga né Western.
- **È errato realizzare *Prenotazione* come relazione tra *Fumetti* e *Clienti* invece di essere una entità come proposto sopra.** Se *Prenotazione* fosse una relazione, non sarebbe possibile per un cliente prenotare lo stesso fumetto (ma in date diverse)

Soluzione “Progettazione Logica”

Il primo passaggio è la ristrutturazione dello ER per rimuovere le generalizzazioni. Di seguito il frammento che viene cambiato.



Alla luce di quanto sopra, lo schema logico è il seguente:

FUMETTO (Titolo, Numero, Anno)

MANGA (Titolo, Numero, Autore)

- Chiave Esterna: MANGA.(Titolo,Numero) → FUMETTO.(Titolo,Numero)

WESTERN (Titolo, Numero, CaseEditrice)

- Chiave Esterna: WESTERN.(Titolo,Numero) → FUMETTO.(Titolo,Numero)

CLIENTE(CF, Nome, Cognome, Email)

COPIA(FumettoTitolo, FumettoNumero, NumeroProgressivo, PrezzoVendita)

- Chiave Esterna: COPIA.(FumettoTitolo, FumettoNumero) → FUMETTO.(Titolo,Numero)

COPIA-USATA(FumettoTitolo, FumettoNumero, NumeroProgressivo, Sconto)

- Chiave Esterna: COPIA-USATA.(FumettoTitolo, FumettoNumero,NumeroProgressivo) → COPIA.(FumettoTitolo, FumettoNumero,NumeroProgressivo)

COMPRA(FumettoTitolo, FumettoNumero, NumeroProgressivo, CF-Cliente, Data, Ora)

- Chiave Esterna: COMPRA.(FumettoTitolo, FumettoNumero,NumeroProgressivo) → COPIA.(FumettoTitolo, FumettoNumero,NumeroProgressivo) →
- Chiave Esterna: COPIA.CF_Cliente → CLIENTE.CF

PRENOTAZIONE(FumettoTitolo, FumettoNumero, CF-Cliente, NumCopie)

- Chiave Esterna: PRENOTAZIONE.(FumettoTitolo, FumettoNumero) → FUMETTO.(Titolo,Numero)
- Chiave Esterna: PRENOTAZIONE.CF-Cliente → CLIENTE.CF

Si noti che, se COMPRA fosse fusa con COPIA, il numero di attributi con valori potenzialmente nulli sarebbero alti. Infatti, COPIA avrebbe tre attributi CF-Cliente, Data e Ora che avrebbero valori nulli per tutte le copie non ancora vendute.

Assegnamento del Punteggio

Parte del Dominio di Modellazione	Modellazione Concettuale					Modellazione Logica (tabelle)					
	Punti per la modellazione	Errori Commessi			Punti per la modellazione	Errori Commessi			1	2	3+
		1	2	3+		1	2	3+			
Modellazione dei fumetti (<i>Entità Fumetto, Manga e Western</i>)	2.1 punti	-0.75 punti	-1.5 punti	-2.1 punti	1.25 punti	-0.4 punti	-0.8 punti	-1.25 punti			
Modellazione delle Copie (<i>Copia, Copia Usata e relazione di</i>)	2.1 punti	-0.75 punti	-1.5 punti	-2.1 punti	1.25 punti	-0.4 punti	-0.8 punti	-1.25 punti			
Modellazione dei clienti e dei suoi acquisti (<i>Compra e Cliente</i>)	1.5 punti	-0.75 punti	-1.25 punti	-1.5 punti	1.5 punti	-0.5 punti	-1 punto	-1.5 punti			
Modellazione delle Prenotazioni (<i>Prenotazione e relazioni collegate</i>) <i>(L'errore di considerare "Prenotazione" come relazione corrisponde di fatto a molteplici errori con una penalizzazione totale di punti: 2.3)</i>	2.3 punti	-0.75 punti	-1.5 punti	-2.3 punti	1.5 punti	-0.5 punti	-1 punto	-1.5 punti			
Conversione allo ER "ristrutturato"	Non applicabile					1.5 punti	-1 punti	-1.5 punto	-1.5 punti		

Esercizio Decomposizione in Terza Forma Normale (6 punti)

Sia data la seguente relazione $R(ABCDE)$, con copertura ridotta
 $G=\{B \rightarrow C, B \rightarrow E, C \rightarrow B \text{ e } C \rightarrow D\}$.

Risolvere i seguenti punti:

- Trovare la/e chiave/i di R , motivando la risposta.
- Effettuare una decomposizione in 3NF ed indicare le chiavi delle relazioni finali ottenute.
- Indicare se la decomposizione ottenuta al punto b è anche in BCNF rispetto all'insieme di dipendenze in G . Motivare la risposta.

Soluzione

Parte a

La chiusura di B è $B^+ = \{B, C, D, E\}$. La chiusura di C è $C^+ = \{B, C, D, E\}$. Ad entrambe le relazioni, manca A . Quindi, AB e AC sono superchiavi ed anche chiavi.

Parte b

Siccome la copertura ridotta è già data come testo dell'esercizio, occorre solamente fare i seguenti passi:

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

Si ottiene un solo insieme $\{B \rightarrow C, B \rightarrow E, C \rightarrow B, C \rightarrow D\}$ perché $B^+ = C^+$.

2. Viene costruita una relazione per ogni sottoinsieme:

$R1(B, C, D, E)$ con due chiavi separate: B e C . Infatti, sappiamo che $B^+ = C^+ = \{B, C, D, E\}$, che sono tutti gli attributi di $R1$.

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

La condizione non si verifica; quindi, rimane solamente la relazione $R1$

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 c'è nessuna relazione che contiene A e B : aggiungiamo $R2(A, B)$ con $\{A, B\}$ chiave

Non c'è nessuna relazione che contiene A e C : aggiungiamo $R3(A, C)$ con $\{A, C\}$ chiave

Parte c

Lo schema con le tre relazioni $R1(B, C, D, E)$, $R2(A, B)$ e $R3(A, C)$ con le chiavi di cui sopra è anche in BCNF. Infatti, le dipendenze funzionali $B \rightarrow C$ e $B \rightarrow E$ rispettano, essendo B chiave di $R1$; inoltre, anche le dipendenze $C \rightarrow B$ e $C \rightarrow D$ rispetto la BCNF

poiché C è chiave di R1. Tali dipendenze non devono essere verificate su R2(A, B) e R3(A, C) poiché:

- $B \rightarrow E$ non va verificata: né R2 né R3 ha sia B che E;
- $B \rightarrow C$ e $C \rightarrow B$ non vanno verificata: né R2 né R3 ha sia B che C;
- $C \rightarrow D$ non va verificata: R2 e R3 non hanno né C né D;

[Assegnamento di punteggio](#)

Gli errori della parte b dovuti alla parte a non vengono penalizzati due volte: si parte dai risultati ottenuti per la parte a. Gli errori della parte c dovuti alla parte b non vengono penalizzati per lo stesso motivo.

- **1.5 punto:** la parte a;
- **2.5 punti:** la parte b;
- **2 punti:** la parte c.



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Seconda Prova del Terzo Appello (data 21 Luglio)

Progettazione Concettuale (8 punti) e Logica (7 punti)

Si richiede di produrre i seguenti artefatti:

- **Schema Concettuale Entità-Relazione**, da inviare per email da account UniPd)
- **Schema Logico**, da inserire nel riquadro di cui sotto. Indicare la chiave primaria, i vincoli di chiave esterne e gli attributi che ammettono valori nulli.

per un'applicazione relativa ad un negozio di noleggio motociclette. Nella progettazione logica, si vuole assicurare il minimo numero minimo di tabelle necessarie, anche a costo di valori nulli. Questo vincolo è al fine di minimizzare il numero di join.

Illustrare come ristrutturare il diagramma ER per essere direttamente traducibile in uno schema relazionale. Il diagramma ER ristrutturato deve anche essere inviato per email da account UniPd.

Il negozio possiede tante motociclette. Di ogni motocicletta, interessa la targa(identificativo), il modello, l'anno di immatricolazione ed i chilometri percorsi. Di ogni modello interessala marca (ad esempio, "Honda"), il codice (unico nell'ambito della marca), la cilindrata, il prezzo del noleggio all'ora ed il prezzo del noleggio al chilometro.

Di ogni noleggio, interessa il cliente che l'ha effettuato, la motocicletta noleggiata, la data in cui è avvenuto (ogni noleggio è infatti relativo ad un giorno), la durata in ore, la tipologia di pagamento (a chilometro oppure ad ora) ed i chilometri percorsi. Si noti che una stessa motocicletta non può essere noleggiata più volte nello stesso giorno.

Dei clienti, Il negozio vuole conoscere il numero di telefono. Ogni cliente possiede una carta di credito, di cui interessa il numero ed il circuito

Per quei noleggi per i quali si è verificato un danno alla motocicletta noleggiata, interessa il codice indicante la causa del danno (se nota) e la persona-conducente che era alla guida della motocicletta al momento del verificarsi del danno. Di ogni conducente, l'applicazione deve memorizzare i seguenti dati: codice fiscale (identificativo), nome, cognome e data di nascita. Si noti che il conducente può coincidere con il cliente del noleggio oppure essere un'altra persona.

(8 punti per la modellazione concettuale + 7 punti per lo schema logico)

Notazione per lo schema logico:

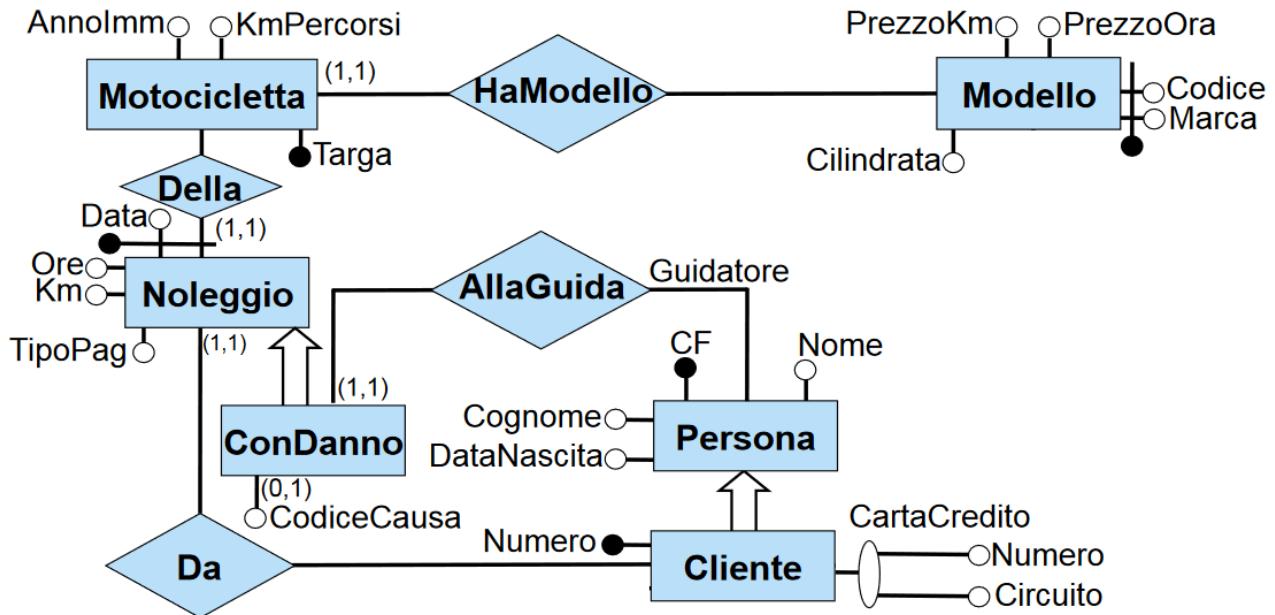
Usare il sottolineato per indicare i vincoli di chiave primaria, e il **bold** per indicare gli attributi che ammettono valori nulli.

Indicare con X.A-->Y.B per indicare che l'attributo A della tabella X è chiave esterna all'attributo B della tabella Y.

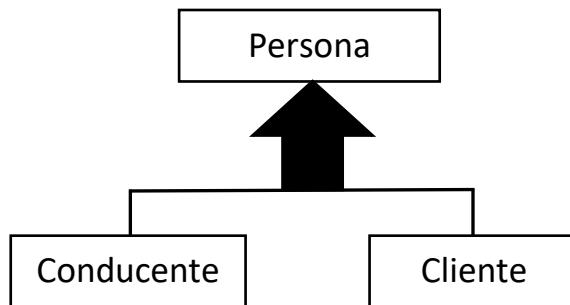
Indicare con X.(A,B) --> Y.(C,D) per indicare che gli attributi (A,B) della tabella X sono chiavi esterna all'attributo (C,D) della tabella Y.

Soluzione Progettazione Concettuale

Una soluzione è data dal seguente schema ER. Alcune cardinalità non sono state specificate per leggibilità del modello (per esempio, per la relazione *HaModello* dal lato del *Modello*). Le cardinalità mancanti nella relazioni sono da intendersi **(0,n)**.¹



Si noti che è errata la seguente soluzione in cui gli attributi di *Persona* rimangono quelli attuali ed, inoltre, *Conducente* è in relazione *AllaGuida* con *ConDanno* e *Cliente* è in relazione con *Noleggio*:

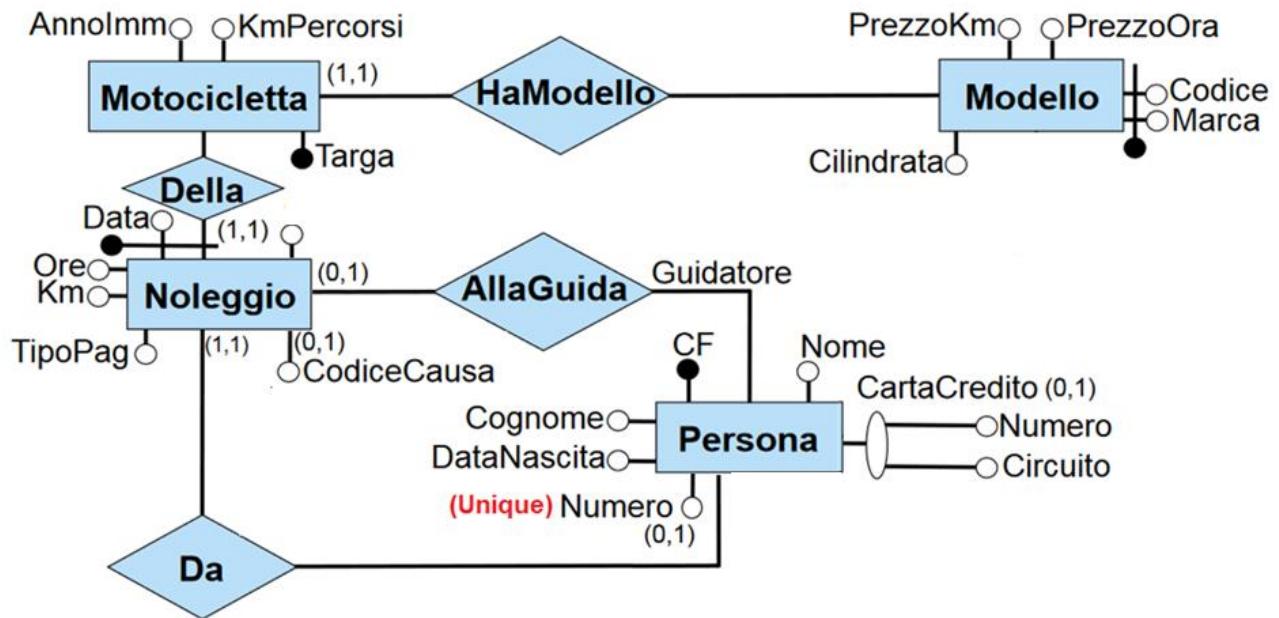


Infatti, questo implicherebbe che un conducente non può essere cliente (si pensi al caso uomo-donna). La generalizzazione parziale o totale (freccia vuota o piena) è irrilevante rispetto all'esclusività di essere conducente o cliente.

¹ Si noti che all'esame occorre inserire tutte le cardinalità: le cardinalità mancanti sono considerate errore.

Soluzione “Progettazione Logica”

Il primo passaggio è la ristrutturazione dello ER per rimuovere le generalizzazioni. Siccome l’obiettivo è minimizzare il numero di tabelle, l’entità *ConDanno* viene fusa con *Noleggio*, l’entità *Cliente* viene fusa con *Persona*. Questo ovviamente introduce un certo numero di attributi con cardinalità (0,1), cioè che ammettono il valore nullo nelle corrispondenti tabelle. Il diagramma ristrutturato è quindi il seguente:



Il vincolo unique per numero serve a rappresentare che, se una persona ha un numero telefonico, allora è l’unica con quel numero. Questo viene derivato dal fatto che l’entità *Cliente* nel diagramma ER concettuale usa tale attributo come identificatore. Si noti che una persona potrebbe non avere un numero di telefono perché solo le persone clienti lo hanno

Assegnamento del Punteggio

La mancanza del vincolo unique non è visto come un errore. La presenza di tale vincolo comporta un bonus di 1 punto.

Parte del Dominio di Modellazione	Modellazione Concettuale					Modellazione Logica (tabelle)				
	Punti per la modellazione	Errori Commessi			Punti per la modellazione	Errori Commessi				
		1	2	3+		1	2	3+		
Modellazione delle Motocicletta e Modello (<i>Entità Modello e Motocicletta e Relazione HaModello</i>)	1.5 punti	-0.5 punti	-1 punto	-1.5 punto	0.75 punti	-0.25 punti	-0.5 punti	-0.75 punti		
Modellazione dei Noleggi (<i>Entità Noleggio e Relazione Della</i>)	1.5 punti	-0.5 punti	-1 punto	-1.5 punto	0.75 punti	-0.25 punti	-0.5 punti	-0.75 punti		
Modellazione dei clienti e la relazione con i noleggi (<i>Da, Cliente e generalizzazione con Persona</i>)	2.5 punti	-0.75 punti	-1.5 punti	-2.5 punti	2 punti	-0.75 punti	-1.5 punti	-2 punti		
Modellazione di Noleggi con Danno (<i>Con Danno, Persona e AllaGuida</i>)	2.5 punti	-0.75 punti	-1.5 punti	-2.5 punti	2 punti	-0.75 punti	-1.5 punti	-2 punti		
Conversione allo ER "ristrutturato"	Non applicabile				1.5 punti	-0.5 punti	-1.5 punti	-1.5 punti		

Esercizio Decomposizione in Terza Forma Normale (6 punti)

Sia data la seguente relazione R(ABCDE), con copertura ridotta $G=\{B \rightarrow C, B \rightarrow E, C \rightarrow A \text{ e } C \rightarrow D\}$.

Risolvere i seguenti punti:

- Trovare la/e chiave/i di R, motivando la risposta.
- Effettuare una decomposizione in 3NF ed indicare le chiavi delle relazioni finali ottenute.
- Indicare se la decomposizione ottenuta al punto b è anche in BCNF rispetto all'insieme di dipendenze in G. Motivare la risposta.

Soluzione

Parte a

La chiusura di B è $B^+ = \{A, B, C, D, E\}$. La chiusura di C è $C^+ = \{A, C, D\}$. Quindi, B è superchiave (ed anche chiave), mentre C non è superchiave poiché mancano B ed E in C^+ .

Parte b

Siccome la copertura ridotta è già data come testo dell'esercizio, occorre solamente fare i seguenti passi:

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

Si ottiene un insieme $\{B \rightarrow C, B \rightarrow E\}$ ed un secondo $\{C \rightarrow A, C \rightarrow D\}$.

2. Viene costruita una relazione per ogni sottoinsieme:

$R1(B, C, E)$ e $R2(C, B, D)$. Siccome ogni relazione ha una sola chiave, questa viene rappresentata sottolineata.

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

La condizione non si verifica; quindi, rimangono $R1$ e $R2$

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):

La condizione non si verifica; quindi, rimangono $R1$ e $R2$

Parte c

Lo schema R1(B, C, E) e R2(C, A, D) è in BCNF:

- $B \rightarrow C$, $B \rightarrow E$ sono le dipendenze relative a R1. La parte di sinistra è B che è superchiave di R1 (anche chiave)
- $C \rightarrow A$, $C \rightarrow D$ sono le dipendenze relative a R2. La parte di sinistra è C che è superchiave di R2 (anche chiave)

Parte d

Gli attributi in comune tra R1 e R2 consistono nel solo C. C è chiave di R2. Siccome condizione sufficiente per una decomposizione senza perdite è che gli attributi in comune siano superchiave di almeno una delle due relazioni, la decomposizione è in effetti senza perdita.

[Assegnamento di punteggio](#)

Siccome le parti b, c, d dipendono in parte dalle parti precedenti, gli errori dovuti ad errori nelle parti che precedono non vengono penalizzati due volte.

Un punteggio di 1.5 viene assegnato ad ogni parte. Per ogni parte, due errori corrispondono ad un punteggio di zero, mentre un primo errore viene penalizzato di 1 punti.



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 22 Gennaio 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 alla gestione delle edicole per la vendita di giornali, Disegnare il diagramma ER nel riquadro della pagina che segue.

Si richiede di progettare lo schema ER concettuale di un'applicazione relativa alle edicole per la vendita di giornali.

Di ogni edicola interessa il comune in cui essa è registrata, il codice, che è unico nell'ambito del comune in cui l'edicola stessa è registrata, la categoria, l'anno di inizio attività (non sempre disponibile), e i contratti che l'edicola ha con i distributori per l'approvvigionamento dei quotidiani.

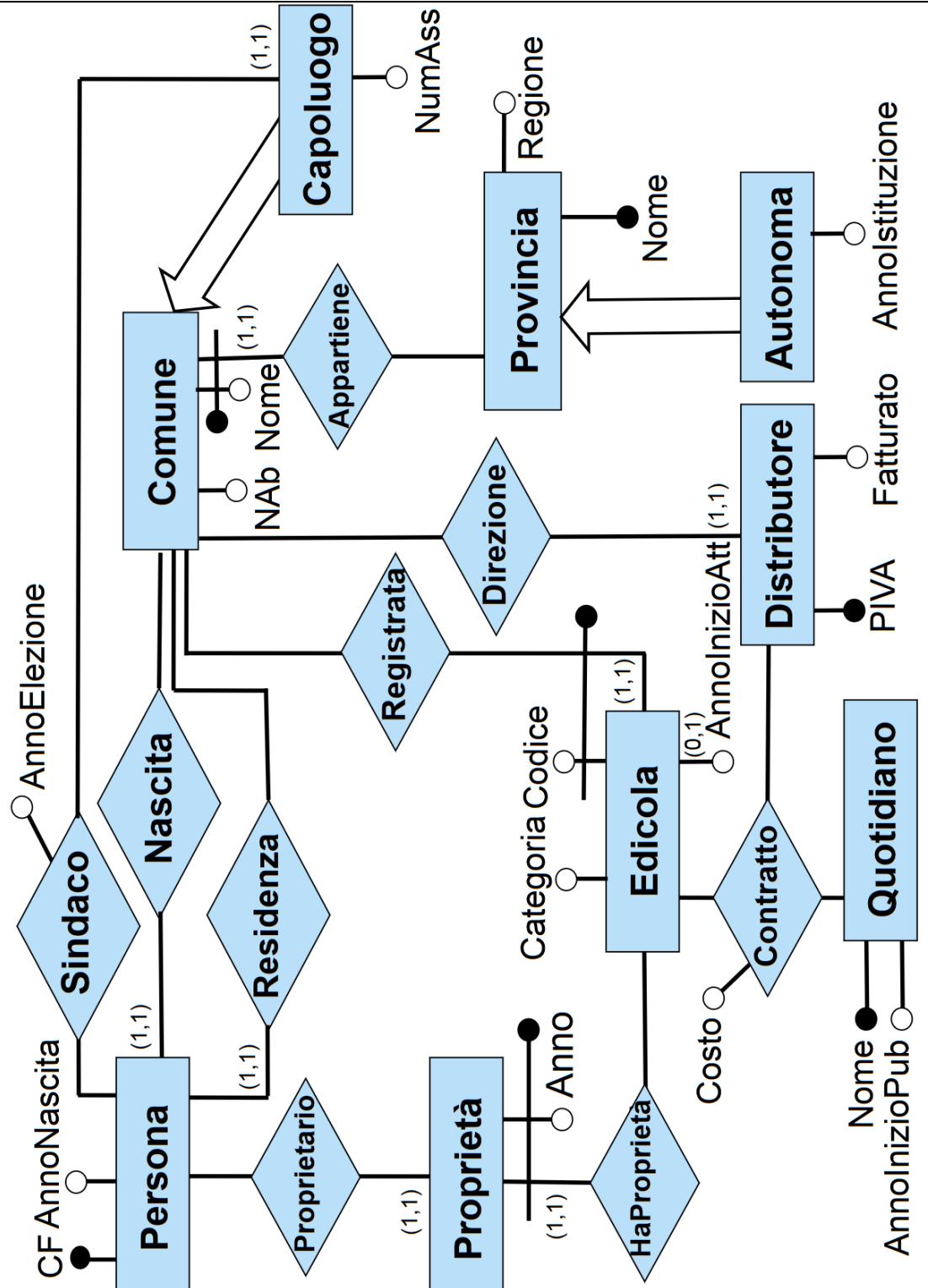
Ogni contratto riguarda un'edicola, un quotidiano ed un distributore, ed è caratterizzato dal costo mensile a carico dell'edicola. Infine, per ogni edicola, interessa conoscere le varie persone che sono state proprietarie dell'edicola nei diversi anni, tenendo conto del fatto che in ogni anno un'edicola ha al massimo un proprietario.

Di ogni persona interessa il codice fiscale (id), l'anno di nascita, il comune di nascita, ed il comune di residenza. Di ogni distributore di quotidiani interessa la partita IVA (id), il fatturato ed il comune in cui è situata la direzione.

Di ogni quotidiano interessa il nome (identificativo), e l'anno di inizio pubblicazione. Di ogni comune interessa la provincia di appartenenza, il nome (unico nella provincia), ed il numero di abitanti. Dei comuni che sono capoluogo di provincia interessa l'attuale sindaco (con l'anno di elezione), ed il numero di assessori comunali.

Di ogni provincia interessa il nome (identificativo) e la regione di appartenenza. Alcune province sono "autonome" e di esse interessa anche l'anno di istituzione.

Soluzione Esercizio 1



Cardinalità (1,n) omesse per leggibilità

Nome: _____ Num. Matric.: _____

Esercizio 2: Progettazione Logica (7 punti)

A partire dallo ER concettuale dell'Esercizio 1, produrre uno schema relazionale del database nel riquadro sottostante, minimizzando i valori nulli. Indicare la chiave primaria, i vincoli di chiave esterne e gli attributi che ammettono valori nulli. Illustrare come ristrutturare il diagramma ER per essere direttamente traducibile in uno schema relazionale.¹

Soluzione (Parziale) Esercizio 2

Circa l'esercizio 1. Non è possibile rappresentare la proprietà come una relazione, poiché una persona può essere proprietaria di una edicola per anni successivi, mentre una relazione permette solamente una tupla (p,e,a) con lo stesso valore di persona *p* e di edicola *e*. D'altronde, una cardinalità (1,N) da entrambi i lati non funzionerebbe neanche, perché permetterebbe a più persone di essere proprietarie di una edicola nello stesso anno.

Ristrutturazione. Volendo minizzare i valori nulli, occorre sostituire le due generalizzazioni con due relazioni con cardinalità (1,1) dal lato di Capolugo e Autonomia e (0,1) dal lato di Comune e Provincia. Gli identificatori di Comune e Autonomia sono le relazioni stesse appena introdotte.

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 con le relazioni:

- **Biblioteca (CodiceBiblio, Citta, Inizio)** che memorizza, per le varie biblioteche, la città in cui si trova, l'anno di inizio delle attività.
 - **Frequentazione (CodiceBiblio, CFPersona)** che memorizza, per ogni codice di biblioteca, il codice fiscale delle persone che hanno visitato quella biblioteca almeno una volta.
- A. Nel riquadro, scrivere una Query in Algebra Relazionale che restituisce le lista dei codici fiscali delle persone che hanno visitato almeno due biblioteche. (2 punti).²

F1=Frequentazione

F2=Frequentazione

$\Pi_{F1.CFPersona}$

$(F1 \bowtie_{F1.CFPersona = F2.CFPersona \text{ AND } F1.CodiceBiblio <> F2.CodiceBiblio} F2)$

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

Nome: _____ Num. Matric.: _____

B. Scrivere una query in Standard SQL per restituire le città in cui tutte le biblioteche che hanno iniziato l'attività nello stesso anno. Per esempio, una citta "X" è restituita, se tutte le biblioteche in "X" hanno iniziato le loro attività nello stesso anno. (2.5 punti).

```
SELECT DISTINCT CITTA  
FROM BIBLIOTECA  
  
EXCEPT  
  
SELECT B1.CITTA  
FROM BIBLIOTECA B1 AND BIBLIOTECA B2  
WHERE B1.CITTA=B2.CITTA AND  
B1.INIZIO<>B2.INIZIO
```

C. Nel riquadro, scrivere una query in Standard SQL che restituisce le biblioteche con un numero di frequentatori più alto della media tra tutte le biblioteche (2.5 punti)

```
CREATE VIEW VISITATORI-PER-BIBLIOTECA  
(CODICEBIBLIO,NUM_VISITATORI) AS  
  
SELECT CODICEBIBLIO, COUNT(*)  
FROM FREQUENTAZIONE  
GROUP BY CODICEBIBLIO;  
  
SELECT CODICEBIBLIO  
FROM VISITATORI-PER-BIBLIOTECA  
WHERE NUM_VISITATORI >  
(SELECT AVG(NUM_VISITATORI)  
FROM VISITATORI-PER-BIBLIOTECA)
```

Esercizio 4: Normalizzazione (5 punti)

Sia data la seguente relazione $R(ABCDE)$, con copertura ridotta $G=\{B \rightarrow C, B \rightarrow E, C \rightarrow B, A \rightarrow B, A \rightarrow D\}$. Risolvere i seguenti punti:

- a. Trovare la/e chiave/i di R , motivando la risposta.
- b. Effettuare una decomposizione in 3NF ed indicare le chiavi delle relazioni finali ottenute.
- c. Indicare se la decomposizione ottenuta al punto b è anche in BCNF rispetto all'insieme di dipendenze in G . Motivare la risposta.

Aggiungere la risposta negli spazi bianchi lasciati appositamente.

Punto a

Si considerano le seguenti chiusure

$$B^+ = \{B, C, E\}$$

$$C^+ = \{B, C, E\}$$

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

A è l'unica chiave.

Punto b

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

Ci sono due gruppi: $\{B \rightarrow C, B \rightarrow E, C \rightarrow B\}$ e $\{A \rightarrow B, A \rightarrow D\}$

2. **Viene costruita una relazione per ogni sottoinsieme:**

$$R1(B, C, E)$$

$$R2(A, B, D)$$

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

La condizione non è verificata, quindi rimangono sia $R1$ che $R2$

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

La condizione non è verificata, quindi rimangono sia $R1$ che $R2$

5. **Indicare le chiavi delle relazioni ottenute dalla normalizzazione**

$$R1(B, C, E) \text{ con chiavi } \{B\} \text{ e } \{C\}$$

$$R2(A, B, D) \text{ con chiave } \{A\}$$

Punto c

La decomposizione è anche il BCNF

Nome: _____ Num. Matric.: _____

Esercizio 5: Quiz (5 punti)

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

Domanda 1 (1 Punto)

Data la query `SELECT * FROM S WHERE X=4 AND Z>8` 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 Hash su (Z,X)
2. Indice B+Tree su X
3. Indice Hash sulla coppia (X,Z)
4. Indice B+Tree sulla coppia (X,Z)

Domanda 2 (1.5 Punti)

Data la relazione $R(A, B, C)$ con una sola dipendenza funzionale $C \rightarrow A$. Una delle seguenti affermazioni è vera. Quale?

1. BC non è né chiave né superchiave
2. BC è chiave ma non è superchiave
3. BC non è chiave ma è superchiave
4. BC è sia chiave che superchiave

Domanda 3 (2 Punti)

Dato il seguente schedule nel log fino ad un guasto

..., CHECKPOINT(T1), BEGIN_TR(T2), ..., BEGIN_TR(T3), ...,
COMMIT(T1), ..., COMMIT (T3), Guasto

Vengono omessi update, insert and delete per leggibilità e perché non rilevanti alla domanda. Di quale/i transazione/i occorre fare il REDO?

1. T2
2. T2 e T1
3. T1 e T3
4. T3



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 12 Febbraio 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 alla gestione delle vigne e delle vendemmie di un produttore di vini. Disegnare il diagramma ER nel riquadro della pagina che segue.

Di ogni vigna interessa il codice (identificatore), la regione in cui si trova, il tipo di uva che vi si coltiva, e i vari produttori che ne sono stati proprietari, con la data in cui lo sono diventati. Si tenga presente che è consentito al più un cambio di proprietà al giorno.

Delle vigne interessano poi anche le vendemmie che sono state effettuate nelle vigne stesse, al massimo una all'anno.

Di ogni vendemmia interessa la vigna in cui è stata eseguita, l'anno in cui è stata eseguita, la valutazione del risultato che ha avuto (un voto da 1 a 10), la persona che ne è stata responsabile (ma non sempre questa informazione è disponibile) e le bottiglie di vino che la vendemmia ha prodotto.

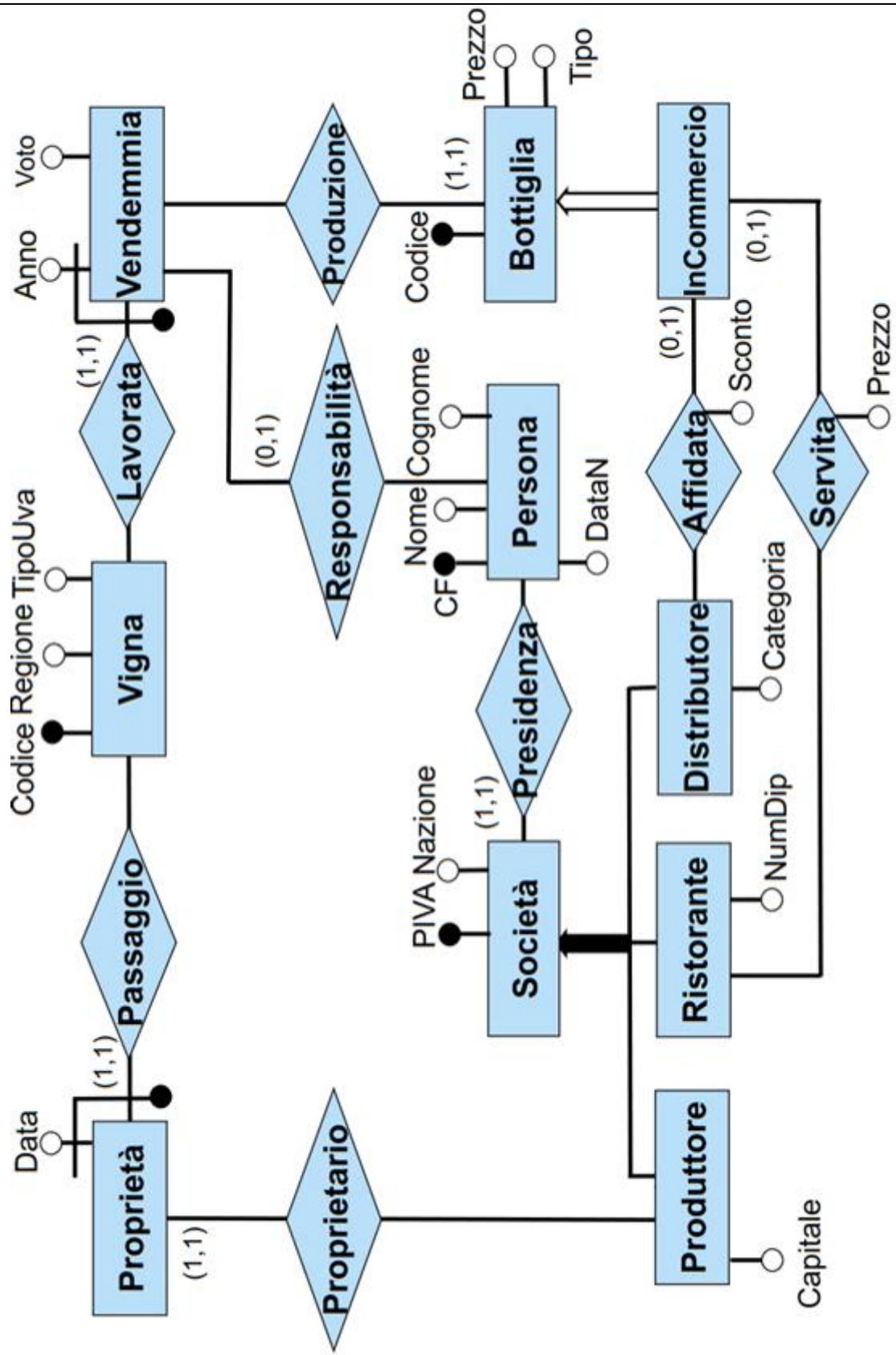
Ogni bottiglia ha un codice identificativo univoco ed è associato con un tipo, un prezzo standard. Non tutte le bottiglie vengono messe in commercio. Quando una bottiglia viene commercializzata, è di interessa sapere l'eventuale distributore che l'ha presa in carico per il suo commercio, con il relativo sconto applicato rispetto al prezzo standard (che può essere anche 0). Ogni bottiglia può essere venduta ad un ristorante: in tal caso, si vuole sapere quale è stato il ristorante e quale sia stato il prezzo di vendita.

Di ogni persona interessa il codice fiscale (identificativo), il nome, il cognome e la data di nascita.

Produttori, distributori e ristoranti sono le società di interesse per la nostra applicazione e di ogni società interessa la partita IVA (identificatore), la nazione in cui è registrata e la persona presidente. In aggiunta, dei produttori interessa il capitale sociale, dei distributori interessa la categoria e dei ristoranti interessa il numero di dipendenti.

Si noti che una società può essere una sola cosa tra produttore, distributore e ristorante.

Soluzione Esercizio 1



Le cardinalità (1,n) omesse per leggibilità

Nome: _____ Num. Matric.: _____

Esercizio 2: Progettazione Logica (7 punti)

A partire dallo 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. Si consideri che, quando si accede al produttore, ristorante o distributore, si vuole sempre sapere la PIVA e la nazione. Tale operazione deve essere ottimizzata. Illustrare come ristrutturare il diagramma ER per essere direttamente traducibile in uno schema relazionale.¹

Soluzione Esercizio 2 (Parziale)

Nella ristrutturazione:

1. L'entità Società viene rimossa e i suoi attributi duplicati per le entità figlie *Produttore, Ristorante e Distributore*;
2. La relazione *Presidenza* viene duplicata per ognuna delle entità figlie.

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 con le relazioni:

- **Programma** (Codice, NomeProgramma, Rete, Tipologia) che, ad ogni codice di programma televisivo, associa il nome del programma, la rete che lo trasmette e la tipologia.
- **Artista** (CProgramma, NomeArtista, compenso) che memorizza gli artisti che hanno partecipato ai diversi programmi, con i rispettivi compensi.

con chiave esterna Artisti.CProgramma → Programmi.Codice.

Si noti che ci possono essere due programmi con lo stesso nome ma codice diverso.

- A. Nel riquadro, scrivere una Query in Algebra Relazionale che restituisce i nomi dei programmi della Rete “Raiuno” in cui si esibiscono almeno due artisti. (2 punti).²

$V1 = \sigma_{\text{Rete}=\text{"Raiuno}} (\text{Programma} \bowtie_{\text{Codice}=\text{CProgramma}} \text{Artista})$

$V1 = V2$

$\Pi_{V1.\text{NomeProgramma}}$

$(V1 \bowtie_{V1.\text{Codice}=V2.\text{Codice} \text{ AND } V1.\text{NomeArtista} \leftrightarrow V2.\text{NomeArtista}} V2)$

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

Nome: _____ Num. Matric.: _____

B. Scrivere una query in Standard SQL che, per ogni codice programma, restituisce il nome del programma e la somma dei compensi per gli artisti del programma. (2.5 punti).

```
CREATE VIEW COMPENO-PROGRAMMA  
(CODICE, COMPENSI) AS  
SELECT CPROGRAMMA, SUM(COMPENO)  
FROM ARTISTA  
GROUP BY CPROGRAMMA;  
  
SELECT C.CODICE, NOMEPROGRAMMA, COMPENSI  
FROM COMPENO-PROGRAMMA, PROGRAMMA  
WHERE COMPENO-  
PROGRAMMA.CODICE=PROGRAMMA.CODICE
```

C. Nel riquadro, scrivere una query in Standard SQL che restituisce la/e rete/i che trasmette più programmi di tipologia “Varietà” (2.5 punti)

```
CREATE VIEW PROGRAMMI-PER-RETE  
(RETE,NUM_PROGRAMMI) AS  
SELECT RETE, COUNT(*)  
FROM PROGRAMMA  
WHERE TIPOLOGIA="Varietà"  
GROUP BY CODICEBIBLIO;  
  
SELECT RETE  
FROM PROGRAMMI-PER-RETE  
WHERE NUM_PROGRAMMI =  
(SELECT MAX(NUM_VISITATORI)  
FROM VISITATORI-PER-BIBLIOTECA)
```

Esercizio 4: Normalizzazione (5 punti)

Sia data la relazione $R(A,B,C,D,E,F)$ e l'insieme di dipendenze associato $\{B \rightarrow A, E \rightarrow D, C \rightarrow B, CE \rightarrow F\}$. Risolvere i seguenti punti:

- a. Trovare la/e chiave/i di R , motivando la risposta.
- b. Indicare quali dipendenze violano la BCNF, motivando la risposta
- c. Effettuare una decomposizione in BCNF senza perdita nel join, indicando le chiavi delle relazioni ottenute
- d. Indicare se la decomposizione ottenuta al punto c preserva le dipendenze. Motivare la risposta.

Punto a

Si considerano le seguenti chiusure

$$\begin{aligned}B^+ &= \{B, A\} \\E^+ &= \{E, D\} \\C^+ &= \{C, B, A\} \\CE^+ &= \{C, E, B, A, F, D\}\end{aligned}$$

Quindi, CE è una chiave della relazione.

Punto b

$B \rightarrow A$, $E \rightarrow D$, $C \rightarrow B$ violano la BCNF perché, rispettivamente, B , E e C non sono superchiavi della relazione R .

Punto c

Usando $B \rightarrow A$

$R1(\underline{B},A)$, rimuovendo A da R che diventa $R(B,C,D,E,F)$

Usando $E \rightarrow D$

$R2(\underline{E},D)$, rimuovendo D da R che diventa $R(B,C,E,F)$

Usando $C \rightarrow B$

$R3(\underline{C},B)$, rimuovendo B da R che diventa $R(C,E,F)$

La relazione R rimanente soddisfa $CE \rightarrow F$: $R(\underline{C},E,F)$

Punto d

Sì, le dipendenze sono preservate poiché, per ogni dipendenza, esiste una relazione che contiene tutti gli attributi

Nome: _____ Num. Matric.: _____

Esercizio 5: Quiz (5 punti)

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

Domanda 1 (1 Punto)

Data la query `SELECT * FROM S WHERE Z=4 ORDER BY X` 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 B+Tree su X
3. Indice Hash sulla coppia (X,Z)
4. Indice B+Tree sulla coppia (Z,X)³

Domanda 2 (2 Punti)

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

1. A volte $0 \leq |R \bowtie_{A=D} S| < |R|$ ed altre volte $|R \bowtie_{A=D} S| = |R|$
2. $0 \leq |R \bowtie_{A=D} S| \leq |R| * |S|$
3. Sempre $|R \bowtie_{A=D} S| = |R|$ ⁴
4. $0 \leq |R \bowtie_{A=D} S| \leq \max(|R|, |S|)$

³ Questo indice permette di trovare la coppia ($Z = 4, X = a$) nell'indice dove a è il valore più piccolo presente per l'attributo X. Si itera sull'indice con coppie ($Z=4, X=a'$) per valori crescenti di X fino a raggiungere la prima coppia con $Z>4$. Vengono restituiti tutte le tuple associate con i valori di indice trovati, che sono già ordinate nel modo giusto

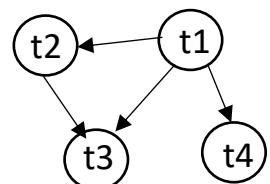
⁴ Essendoci la chiave esterna specificata, ogni tupla di R è tale che l'attributo A è tra i valori dell'attributo S di una tupla di D. Quindi, ogni tupla di R rimane del prodotto cartesiano. Inoltre, poiché D è chiave primaria di S, ogni tupla di R è associata ad esattamente una tupla di S. Quindi, il numero di tuple del risultato è esattamente uguale alle tuple di R. Si noti che non è vero il contrario: ci sono tuple di S che non sono associate a tuple di R.

Domanda 3 (2 Punti)

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

W₁(A) R₂(A) R₂(B) W₂(D) R₃(C) R₁(C) W₃(B) R₄(A) 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 ⁵
4. S **non** è né conflict-serializzabile né view-serializzabile

⁵ Il grafo dei conflitti è aciclico, il ché implica la conflict-serializzabilità, la quale implica a sua volta la view-serializzabilità



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.

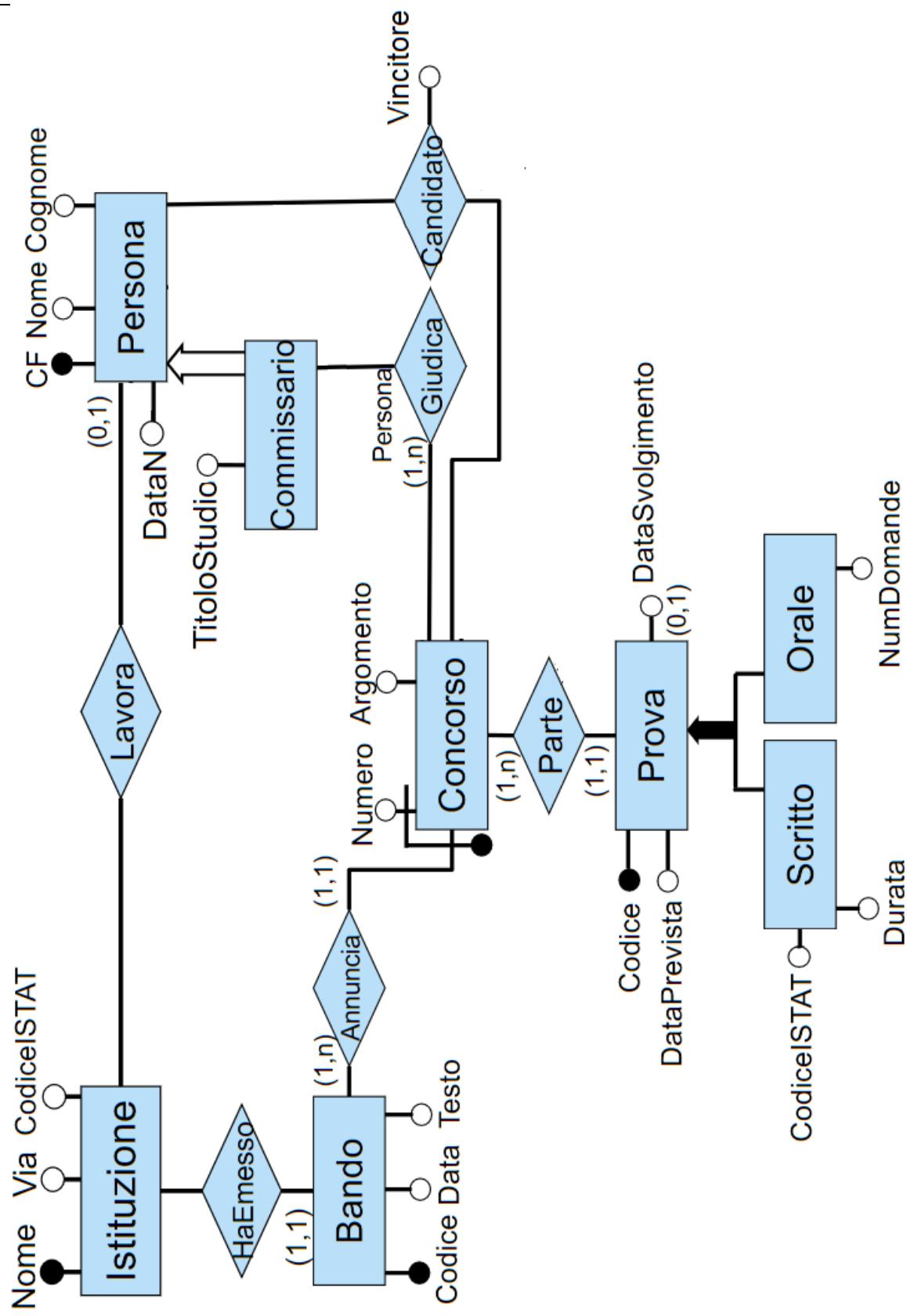
Di ogni bando interessa il codice identificativo, la data di pubblicazione, il testo ed i concorsi (almeno uno) che sono annunciati nel bando stesso.

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), indentificata 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 (Città, Servizio, Anno, Spesa)**
- **POSIZIONE (Città, 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 =  
ΠS1.CITTA (S1 ⋈S1.CITTA=S2.CITTA AND S1.SERVIZIO<>S2.SERVIZIO S2)  
  
ΠCITTA (S1) \ CITTA_DUE
```

² Si assuma che l'operatore di join A⋈B **senza condizioni** mantenga le tuple di A X B con valori uguali su attributi uguali (join naturale). Se **una condizione C è specificata**, ⋈_C mantiene le tuple di A X 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. Siene costruita una relazione per ogni sottoinsieme:

R1 (C, B, D, F, E), R2(B, A), R3(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

R1(C, B, D, F, E) con chiave C (aggiungendo anche chiave BE , si ottiene una decomposizione BCNF).

R2(B, A) con chiave B

R3(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 delle relazioni $R(A, B, C)$ e $S(D, E, F)$ 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 $R \bowtie_{A=E} S$?

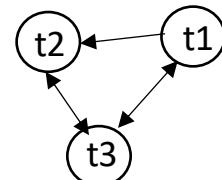
1. $0 \leq |R \bowtie_{A=E} S| \leq |R|$
2. $0 \leq |R \bowtie_{A=E} S| \leq |R| * |S|$
3. $|R \bowtie_{A=E} S| = |S|$
4. $0 \leq |R \bowtie_{A=E} S| \leq \max(|R|, |S|)$

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

Nome: _____ Num. Matric.: _____



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 21 Settembre 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'esame
 - Riceveranno 0 punti "di ufficio"
 - Dovranno saltare l'appello successivo

Codice Easybadge: 420884

Esercizio 1: Diagramma ER (8 punti)

*Si richiede di produrre lo schema concettuale Entità-Relazione di un'applicazione relativa ai festival musicali organizzati da agenzie.
Disegnare il diagramma ER nel riquadro della pagina che segue.*

Si richiede di progettare lo schema concettuale Entità-Relazione di un'applicazione relativa ai festival musicali organizzati da agenzie.

Di ogni agenzia interessa il nome (identificativo), la provincia della sede (con codice ISTAT identificativo e regione), la persona che la dirige (con anno di inizio direzione) e gli avvisi di festival che ha eventualmente pubblicato.

Festival sono annunciati tramite avvisi. Di ogni avviso interessa l'agenzia che lo ha pubblicato, il codice (unico nell'ambito dell'agenzia che lo ha pubblicato), la data di pubblicazione ed il testo definitivo.

Gli avvisi si suddividono in annullati e validi. Dei primi interessa il motivo dell'annullamento e dei secondi interessa il festival che è annunciato nell'avviso.

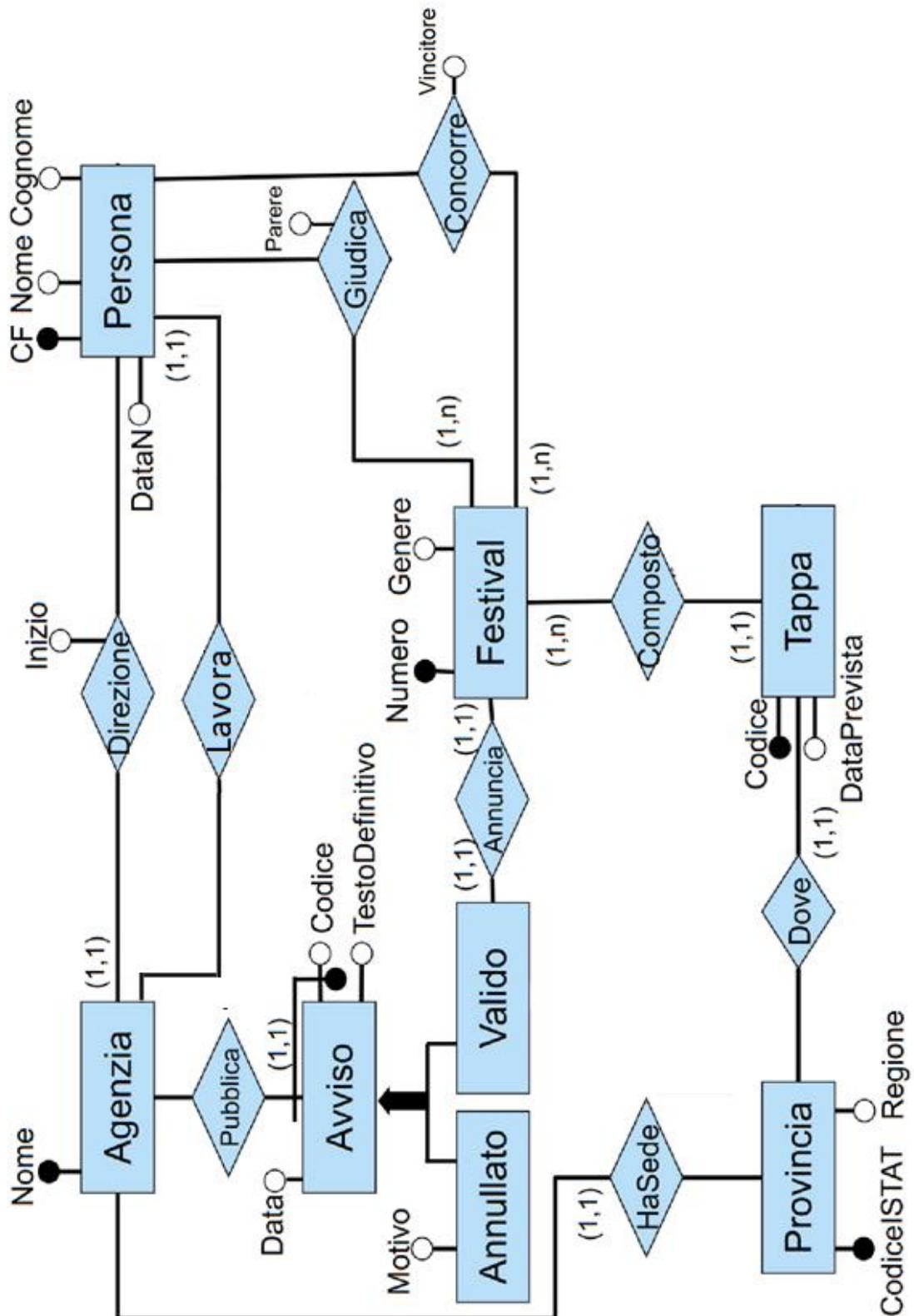
Di ogni festival interessa l'avviso in cui è annunciato, il numero (identificativo), il genere di musica coperto (ad esempio, classica, pop, rock, ecc.) e le tappe di cui è composto (almeno una), ciascuna con codice identificativo (univoco in assoluto), data di svolgimento prevista e la provincia in cui svolgerà.

Di ogni festival interessa sapere le tappe (almeno una), le persone che hanno partecipato come giurati (almeno una), le persone che hanno partecipato come concorrenti (almeno una) e, tra queste ultime, quelle che sono risultate vincitrici (se ci sono).

Di ogni persona interessa il codice fiscale (identificativo), il nome, il cognome e la data di nascita. ~~Infine, di ogni persona che ha svolto il ruolo di giurato interessa anche l'eventuale agenzia in cui lavora.~~

Nome: _____ Num. Matric.: _____

Soluzione Esercizio 1



Le cardinalità (0,n) vengono omesse per leggibilità

Esercizio 2: Progettazione Logica (7 punti)

A partire dal Diagramma ER dell'Esercizio 1, produrre uno schema relazionale del database nel riquadro sottostante, minimizzando i valori nulli delle relazioni. Indicare i vincoli di chiave e gli attributi che ammettono valori nulli. Illustrare come ristrutturare l'ER per essere traducibile in uno schema relazionale.¹

Soluzione Esercizio 2

¹ Per mostrare la ristrutturazione dell'ER, è possibile ridisegnare il diagramma o mostrare i cambiamenti apportati al diagramma nel riquadro soluzione dell'Esercizio 1, con una penna di diverso colore (ma non rossa!)

Nome: _____ Num. Matric.: _____

Esercizio 3: Algebra Relazionale & SQL (7 punti)

Si consideri la seguente base di dati per la registrazione degli impiegati che lavorano in aziende:

- IMPIEGATO(CF, Nome, Cognome, Residenza)
- LAVORA(CF, PIVA, DataInizio, DataFine, StipendioMensile)
- AZIENDA(PIVA, Citta, Regione)

dove nessun attributo ammette valori nulli tranne DataFine che può essere NULL se l'impiegato lavora ancora in una azienda. Si noti che una persona può lavorare contemporaneamente in più aziende.

- A. Nel riquadro, scrivere una query in algebra relazione che restituisca il codice fiscale delle persone che attualmente hanno esattamente un solo lavoro (2 punti).²

$L1 = \sigma_{DATAFINE \text{ IS NULL}} (LAVORA)$

$L2 = L1$

$\pi_{L1.CF} (L1) -$

$\pi_{L1.CF} (L1 \bowtie L1.CF = L2.CF \text{ AND } L1.PIVA} \bowtie L2.PIVA L2)$

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.

B. Nel riquadro, scrivere una query in Standard SQL che restituisce la partita IVA dell'azienda con più dipendenti impiegati alla data attuale. (2.5 punti).

```
CREATE VIEW NUM_DIPENDENTI(PIVA,NUM_DIP)
SELECT PIVA, COUNT(*)
FROM LAVORA
WHERE DATAFINE IS NULL
GROUP BY PIVA

SELECT PIVA
FROM NUM_DIPENDENTI
WHERE NUM_DIP =
    (SELECT MAX(NUM_DIP) FROM NUM_DIPENDENTI)
```

C. Nel riquadro, scrivere una query in Standard SQL che, per ogni partita IVA in Veneto relativo ad aziende con almeno 10 dipendenti, restituisce lo stipendio medio (2.5 punti)

```
SELECT PIVA, AVG(STIPENDIOMENSILE)
FROM LAVORA
WHERE DATAFINE IS NULL AND PIVA IN
    (SELECT PIVA FROM AZIENDA WHERE REGIONE='Veneto')
GROUP BY PIVA
HAVING COUNT(CF)>9
```

La parte in grassetto serve ad assicurare che il calcolo sia per i dipendenti attuali: non è strettamente richiesto.

Nome: _____ Num. Matric.: _____

Esercizio 4: Normalizzazione (5 punti)

Sia data la seguente relazione R(ABCDE), con copertura ridotta
 $G=\{B \rightarrow C, B \rightarrow E, C \rightarrow A, CD \rightarrow E\}$.

- a. Trovare la/e chiave/i di R, motivando la risposta.
- b. Effettuare una decomposizione in 3NF ed indicare le chiavi delle relazioni finali ottenute.
- c. Indicare se la decomposizione ottenuta al punto b è anche in BCNF rispetto all'insieme di dipendenze in G. Motivare la risposta.
- d. Indicare se c'è conservazione delle dipendenze. Motivare la risposta.

Parte a

La chiusura di B è $B^+ = \{A, B, C, E\}$. La chiusura di C è $C^+ = \{C, A\}$. La chiusura di CD è $CD^+ = \{A, C, D, E\}$

Quindi, B non è superchiave. Ma aggiungendo D, BD^+ contiene tutti gli attributi e quindi BD è chiave.

Parte b

Siccome la copertura ridotta è già data come testo dell'esercizio, occorre solamente fare i seguenti passi:

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

Si ottiene un insieme $\{B \rightarrow C, B \rightarrow E\}$, un secondo $\{C \rightarrow A\}$ e un terzo $\{CD \rightarrow E\}$.

2. Viene costruita una relazione per ogni sottoinsieme:

R1(B, C, E), R2(C, A) e R3 (C, D, E). Siccome ogni relazione ha una sola chiave, questa viene rappresentata sottolineata.

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

La condizione non si verifica; quindi, rimangono R1, R2 e R3.

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):

Aggiungo relazione R4(B,D).

Parte c

La decomposizione è anche in BCNF. Le dipendenze $B \rightarrow C$, $B \rightarrow E$ si applicano a R1 perché contengono B,C e E: B è chiave di R1. La dipendenza $C \rightarrow A$ si applica a R2 per cui C è chiave. La dipendenza $CD \rightarrow E$ si applica a R3 per cui CD è chiave.

Parte d

Si: per ogni dipendenza funzionale, c'è una relazione che contiene tutti gli attributi della dipendenza: gli attributi di $B \rightarrow C$, $B \rightarrow E$ sono in R1, gli attributi di $C \rightarrow A$ sono in R2 e quelli di $CD \rightarrow E$ sono in R3.

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 relazione R(A, B, C) con dipendenze funzionali $A \rightarrow BC$ e $B \rightarrow A$. Una delle seguenti affermazioni è vera. Quale?

- (1) AB è superchiave ma non è chiave;
- (2) AB è sia chiave che superchiave;
- (3) AB non è né chiave né superchiave;
- (4) AB è chiave ma non è superchiave.

Domanda 2 (1.5 Punti)

Sia data la seguente porzione di log fino al guasto: CK(T5,T6), B(T7), U(T7,O6,B6,A6), U(T6, O3, B7, A7), B(T8), C(T7), I(T8,O5,A5). Quali transazioni richiedono l'UNDO?

1. T5,T6,T8
2. T5,T6,T7,T8
3. T7
4. T7,T8

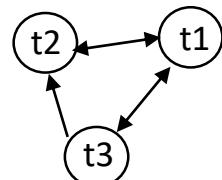
Domanda 3 (2 Punti)

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

$$S = r1(x), r2(y), r3(x), w3(x), w1(x), w1(y), r2(x), w2(x)$$

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





Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 22 Gennaio 2022

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'esame**
 - **Riceveranno 0 punti “di ufficio”**
 - **Dovranno saltare l'appello successivo**

Codice Easybadge: 951112

Esercizio 1: Diagramma ER (8 punti)

Si richiede di produrre lo schema concettuale Entità-Relazione di un'applicazione relativa ad un'azienda che fabbrica e vende mobili. Disegnare il diagramma ER nel riquadro della pagina che segue.

Di ogni mobile interessa il “codice unico mobile” (CUM), che identifica il mobile, il numero di giorni impiegati per la sua lavorazione e le parti utilizzate per la sua costruzione. Infatti, ogni mobile è costruito assemblando 2 o più parti e di ogni parte interessa il codice (identificativo), il tipo, ed il tronco usato per produrla (uno ed uno solo).

L'applicazione memorizza le informazioni di vari luoghi geografici (vedi sotto). Di tutti i luoghi geografici interessa il codice (identificativo) e l'area in chilometri quadrati.

Di ogni mobile interessa sapere chi sono i clienti che lo hanno prenotato ed in quale data lo hanno prenotato. Tra i clienti che hanno prenotato un mobile l'azienda sceglie il cliente al quale vendere il mobile stesso, insieme al prezzo di vendita. Di ogni cliente interessa il codice fiscale (identificativo), la data di nascita e la città di residenza.

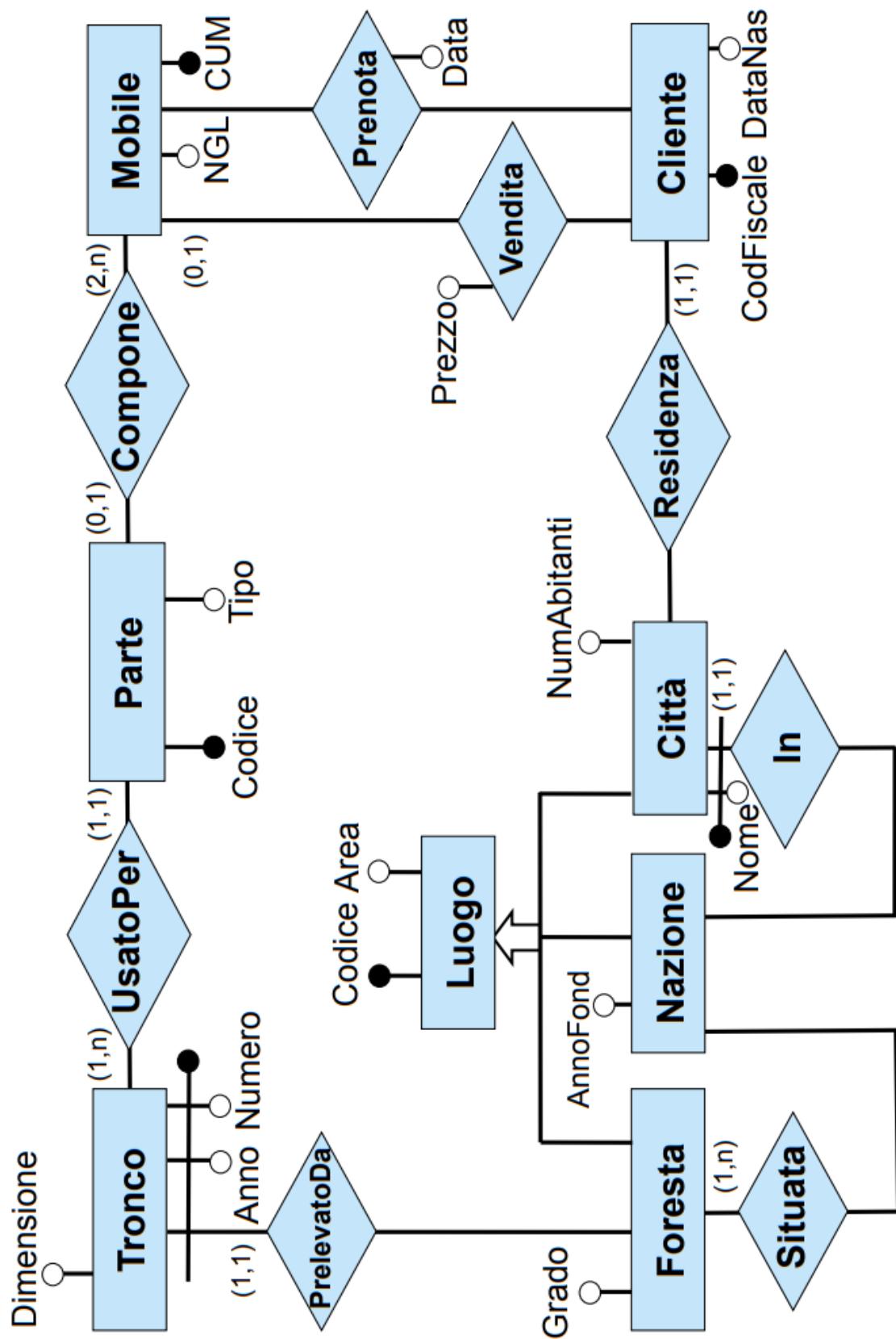
Le città sono luoghi geografici di cui, oltre alle proprietà di tutti i luoghi geografici, interessa il numero di abitanti, la nazione in cui si trova ed il nome (unico nell'ambito della nazione in cui si trova).

I tronchi che interessano all'applicazione sono quelli usati per produrre almeno una parte. Ogni tronco viene prelevato da una foresta, ed ha una dimensione. Ad ogni tronco prelevato viene assegnato, usando un contatore, un numero progressivo unico nell'ambito della foresta (all'inizio di ogni anno, il contatore associato ad ogni foresta viene azzerato).

Di ogni foresta, oltre alle proprietà di tutti i luoghi geografici, interessa il grado di inquinamento e le nazioni (almeno una) in cui è situato il suo territorio. Di ogni nazione, oltre alle proprietà di tutti i luoghi geografici, interessa l'anno della sua fondazione.

Nome: _____ Num. Matric.: _____

Soluzione Esercizio 1



Le cardinalità $(0,n)$ vengono omesse per leggibilità. È ammesso che la generalizzazione sia totale

Esercizio 2: Progettazione Logica (7 punti)

A partire dal Diagramma ER dell'Esercizio 1, produrre uno schema relazionale del database nel riquadro sottostante, minimizzando i valori nulli delle relazioni. Indicare i vincoli di chiave e gli attributi che ammettono valori nulli. Illustrare come ristrutturare l'ER per essere traducibile in uno schema relazionale.¹

Soluzione Esercizio 2

¹ Per mostrare la ristrutturazione dell'ER, è possibile ridisegnare il diagramma o mostrare i cambiamenti apportati al diagramma nel riquadro soluzione dell'Esercizio 1, con una penna di diverso colore (ma non rossa!)

Nome: _____ Num. Matric.: _____

Esercizio 3: Algebra Relazionale & SQL (7 punti)

Si consideri la seguente base di dati per la registrazione dei voli:

- AEROPORTO(Città, Nazione, NumPiste)
- VOLO(IdVolo, GiornoSett, CittaPart, OraPart, CittàArr, OraArr, TipoAereo)

A. Nel riquadro, scrivere una query in Algebra Relazionale che restituisce la/e città da cui partono solo voli internazionali (2 punti).²

AV_P =

(AEREOPORTO \bowtie CITTAPART=CITTA VOLO)

AV_A =

(AEREOPORTO \bowtie CITTAARR=CITTA VOLO)

$\Pi_{CITTAPART} (AV_P) -$

$\Pi_{AV_P.CITTAPART}$

$(AV_P \bowtie_{AV_P.IDVOLO=AV_A.IDVOLO \text{ AND}} AV_P.NAZIONE=AV_A.NAZIONE AV_A)$

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

B. Nel riquadro, scrivere una query SQL che restitusce il tipo di aereo più utilizzato in arrivo nell'aeroporto della città di Roma il giovedì (2.5 punti).

```
CREATE VIEW NUM_VOLI(TIPO,NUMERO)
SELECT TIPOAEREO,COUNT(*)
FROM VOLO
WHERE CITTAARR='Roma' AND GIORNOSETT='Giovedì'
GROUP BY TIPOAEREO

SELECT PIVA
FROM NUM_VOLI
WHERE NUMERO =
      (SELECT MAX(NUMERO) FROM NUM_VOLI)
```

C. Nel riquadro, scrivere una query SQL che restitusce l'id e la città di arrivo dell'ultimo volo in partenza il giovedì dall'aeroporto di Napoli (2.5 punti).

```
SELECT IDVOLO,CITTAARR
FROM VOLO
WHERE CITTAPART='Napoli' AND GIORNOSETT='Giovedì'
AND ORAPART=
      (SELECT MAX(ORAPART)
       FROM VOLO
       WHERE CITTAPART='Napoli' AND GIORNOSETT='Giovedì')
```

Nome: _____ Num. Matric.: _____

Esercizio 4: Normalizzazione (5 punti)

Sia data la seguente relazione $R(ABCDE)$, con copertura ridotta $G=\{B \rightarrow C, B \rightarrow E, C \rightarrow A \text{ e } C \rightarrow D\}$.

Risolvere i seguenti punti:

- Trovare la/e chiave/i di R , motivando la risposta.
- Effettuare una decomposizione in 3NF ed indicare le chiavi delle relazioni finali ottenute.
- Indicare se la decomposizione ottenuta al punto b è anche in BCNF rispetto all'insieme di dipendenze in G . Motivare la risposta.

Soluzione

Parte a

La chiusura di B è $B^+ = \{A, B, C, D, E\}$. La chiusura di C è $C^+ = \{A, C, D\}$. Quindi, B è superchiave (ed anche chiave), mentre C non è superchiave poiché mancano B ed E in C^+ .

Parte b

Siccome la copertura ridotta è già data come testo dell'esercizio, occorre solamente fare i seguenti passi:

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

Si ottiene un insieme $\{B \rightarrow C, B \rightarrow E\}$ ed un secondo $\{C \rightarrow A, C \rightarrow D\}$.

2. Viene costruita una relazione per ogni sottoinsieme:

$R1(\underline{B}, C, E)$ e $R2(\underline{C}, A, D)$. Siccome ogni relazione ha una sola chiave, questa viene rappresentata sottolineata.

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

La condizione non si verifica; quindi, rimangono $R1$ e $R2$

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

La condizione non si verifica; quindi, rimangono $R1$ e $R2$

Parte c

Lo schema $R1(\underline{B}, C, E)$ e $R2(\underline{C}, A, D)$ è in BCNF:

- $B \rightarrow C, B \rightarrow E$ sono le dipendenze relative a $R1$. La parte di sinistra è B che è superchiave di $R1$ (anche chiave)
- $C \rightarrow A, C \rightarrow D$ sono le dipendenze relative a $R2$. La parte di sinistra è C che è superchiave di $R2$ (anche chiave)

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 relazione $R(A, B, C, D)$, indicato con $|R|$ il numero di tuple di R .

Quante tuple sono presenti nel risultato della seguente operazione in Algebra Relazionale $\sigma_{A='val1' \text{ AND } B='val2'}(R)$?

- (1) Il numero di tuple è sempre minore di 2;
- (2) Il numero può essere sia minore di $|R|$, che uguale a $|R|$;
- (3) Il numero è sempre uguale a $|R|$
- (4) Il numero è sempre uguale a 0

Domanda 2 (1.5 Punti)

Sia data la seguente porzione di log fino al guasto: CK(T5,T6), B(T7), U(T7,O6,B6,A6), U(T6, O3, B7, A7), B(T8), C(T7), I(T8,O5,A5). Quali transazioni richiedono il REDO?

1. T5,T6,T8
2. T5,T6,T7,T8
3. T7
4. T7,T8

Domanda 3 (2 Punti)

Si consideri le relazioni $R(A, B, C, D)$ e la seguente query

SELECT MIN(A) FROM R WHERE B=10

Quale dei seguenti indici garantisce l'efficienza massima?

1. Indice Hash sulla coppia (B,A)
2. Indice Hash sulla coppia (A,B)
3. Indice B-TREE sulla coppia (A,B)
4. Indice B-TREE sulla coppia (B,A)



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 14 Giugno 2022

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 (6.5 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.

Si richiede di progettare lo schema concettuale Entità-Relazione di un'applicazione relativa ad una compagnia operante nel campo della cybersecurity, che raccoglie dati su malware, antimalware ed utenti.

Di ogni malware¹ interessa il codice (identificatore), l'anno in cui è apparso ed il nome dei sistemi operativi (almeno uno) in cui può operare. Di ogni antimalware interessa il codice (identificativo), il costo di sviluppo ed i malware (almeno uno) che esso è potenzialmente in grado di rilevare.

Di ogni utente interessa il codice fiscale (identificativo), l'anno di registrazione e gli episodi di infezione di cui è stato vittima, ciascuno con le seguenti caratteristiche: (i) il malware coinvolto, (ii) il danno monetario subito e (iii) la data in cui l'infezione è avvenuta. Si noti che un utente può infettarsi dello stesso malware, ma in giorni diversi.

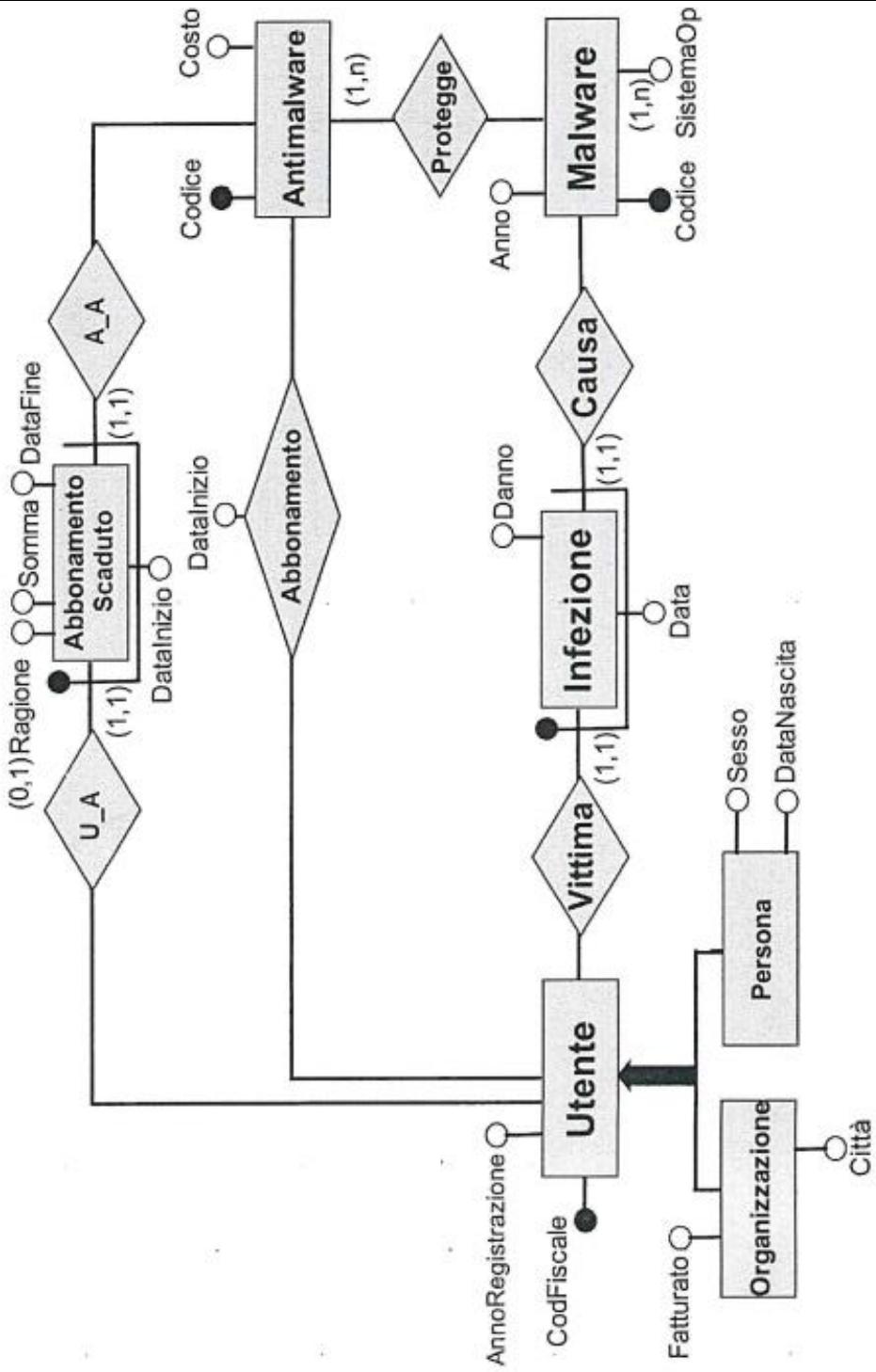
Di ogni utente interessa anche sapere quali sono stati gli antimalware di cui si è dotato nel tempo mediante abbonamento. In particolare:

- riguardo agli antimalware ai quali l'utente è attualmente abbonato interessa sapere la data di inizio dell'abbonamento.
- riguardo agli abbonamenti scaduti di antimalware, interessa la somma complessivamente pagata dall'utente per l'abbonamento, la data di inizio e la data di fine dello stesso e la ragione della disdetta (se tale ragione è nota).

Per utente nel nostro contesto si intende o una persona o una organizzazione di interesse per la compagnia di cybersecurity. Di ogni persona interessa il sesso, la data di nascita. Di ogni organizzazione interessa il fatturato e la città in cui è ubicata la sede legale

¹ Un malware è un programma maligno il cui scopo è danneggiare i computer in cui riesce ad insinuarsi

Soluzione Esercizio 1

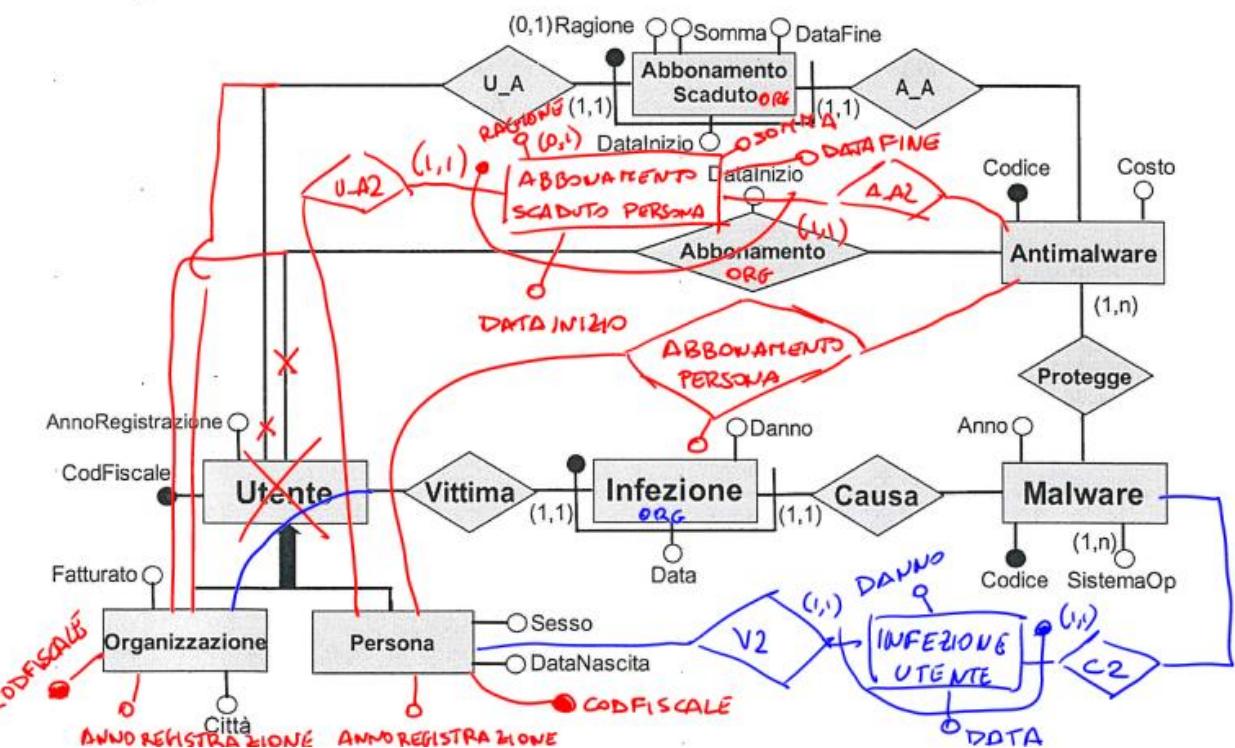


Le cardinalità (0,n) omesse per leggibilità. Si osservi come Infezione e Abbonamento Scaduto siano implementate come entità e non relazione. Se fossero relazioni, (1) non sarebbe possibile per un utente di essere infettato dallo stesso malware in date diverse (si noti come Data sia parte dell'identificatore), (2) non sarebbe possibile per un utente di avere avuto abbonamenti in passato per lo stesso malware ma con data inizio diversa.

Esercizio 2: Progettazione Logica (7.5 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 velocizzare la seguente operazione di join senza introdurre valori nulli: se si accede ad un utente, si accede anche alle informazioni sull'organizzazione o sulla persona.** Illustrare come ristrutturare l'ER per essere traducibile in uno schema relazionale.²

Soluzione Esercizio 2 (Parziale)



Per ottimizzare l'accesso congiunto alle informazioni di un utente e dell'organizzazione o della persona, occorre incorporare *Utente* in *Organizzazione* e *Persona*. Questo richiede di duplicare le entità *AbbonamentoScaduto*, e *Infezione*, e la relazione *Abbonamento*, come mostrato in figura. I cambiamenti vengono intenzionalmente mostrati con una penna per illustrare come sia possibile mostrare la ristrutturazione in sede di esame, cioè sul diagramma originario usando una penna di diverso colore (non rossa, anche se qui viene mostrato in rosso per maggiore leggibilità!).

Si noti che la soluzione di incorporare le entità figlie nel padre *Utente* avrebbe anche l'effetto di ottimizzare l'accesso congiunto, ma causerebbe l'introduzione di parecchi valori nulli.

² 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 per le votazioni locali in diverse città italiane:

- **RISULTATO-PARTITO** (Città, Partito, Anno, Percentuale)
- **CANDIDATURA** (Candidato, Anno, Partito, Città, NumVoti)

Una tupla (PD, Padova, 2022, 30) in RISULTATO-PARTITO indica che il partito PD ha preso il 30% dei voti a Padova nel 2022. Una tupla (Giordani,2022,PD,Padova,3000) in CANDIDATURA indica che Giordani ha preso 3000 voti a Padova nel 2022 con il PD.

- A. Restituire il nome del candidato che si sono presentati una sola volta (cioè in un solo anno) per il PD (2 punti).³

$C1 = \sigma_{\text{Partito}='PD'} (\text{CANDIDATURA})$

$C2 = C1$

$\text{DUE_CANDIDATURE} =$

$\Pi_{C1.\text{CANDIDATO}}$

$(C1 \bowtie_{C1.\text{CANDIDATO}=C2.\text{CANDIDATO} \text{ AND } C1.\text{PARTITO} <> C2.\text{PARTITO}} C2)$

$(\Pi_{C1.\text{CANDIDATO}} (C1)) \setminus \text{DUE_CANDIDATURE}$

³ 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. Nel riquadro, scrivere una query in Standard SQL che restituisce, per ogni partito che ha ottenuto almeno il 10% a Padova nel 2022, il numero medio di voti tra tutti i candidati del 2022. (2.5 punti). ⁴

```
SELECT PARTITO, AVG(NUMVOTI)
FROM CANDIDATURA C
WHERE ANNO=2022 AND CITTA='Padova'
AND PARTITO IN
  (SELECT PARTITO FROM RISULTATO-PARTITO
   WHERE PERCENTUALE>=10 AND ANNO=2022
   AND CITTA='Padova')
GROUP BY PARTITO
```

C. Nel riquadro, scrivere una query in Standard SQL che, per ogni città con votazioni nel 2022, restituisce il partito che ha ottenuto più voti (2.5 punti)

```
CREATE VIEW TOP-PERCENTUALE-
CITTA(CITTA, TOP) AS
  SELECT CITTA, MAX(PERCENTUALE)
    FROM RISULTATO-PARTITO
   WHERE ANNO=2022
 GROUP BY CITTA

  SELECT CITTA, PARTITO
    FROM RISULTATO-PARTITO R,
         TOP-PERCENTUALE-CITTA T
   WHERE R.CITTA=T.CITTA
```

⁴ Vengono quindi restituite coppie del tipo (PD,50) ad indicare che i candidati del PD hanno ottenuto in media 50 voti.

Esercizio 4: Normalizzazione (7.5 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:

R1 (C, B, D, F, E), R2(B, A), R3(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

R1(C, B, D, F, E) con chiave C (aggiungendo anche chiave BE , si ottiene una decomposizione BCNF).

R2(B, A) con chiave B

R3(A, D, E) con chiave AD

Nome: _____ Num. Matric.: _____

Esercizio 5: Quiz (3.5 punti)

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

Domanda 1 (1.5 Punti)

Sia data la seguente porzione di log fino al guasto: CK(T5,T6), B(T7), U(T7,O6,B6,A6), U(T6, O3, B7, A7), B(T8), I(T8,O5,A5), C(T8), A(T5). Quali transazioni richiedono il REDO?

1. T5,T6,T8
2. T5,T6,T7,T8
3. T8
4. T7,T8

Domanda 3 (2 Punti)

Si consideri le relazioni R(A, B, C, D) e la seguente query

SELECT MIN(B) FROM R WHERE A=10

Quale dei seguenti indici garantisce l'efficienza massima?

1. Indice Hash sulla coppia (B,A)
2. Indice Hash sulla coppia (A,B)
3. Indice B-TREE sulla coppia (A,B)
4. Indice B-TREE sulla coppia (B,A)



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 1 Luglio 2022

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 a ricercatori biologi e i loro esperimenti. Disegnare il diagramma ER nel riquadro della pagina che segue.

Di ogni ricercatore interessa il codice fiscale (identificativo), la data di nascita, la nazione di nascita, gli esperimenti effettuati ed i laboratori in cui ha lavorato (almeno uno), con data di inizio e data di fine rapporto (ovviamente, quest'ultima informazione manca per il laboratorio in cui il ricercatore lavora attualmente) e stipendio medio ottenuto nel periodo del rapporto stesso.

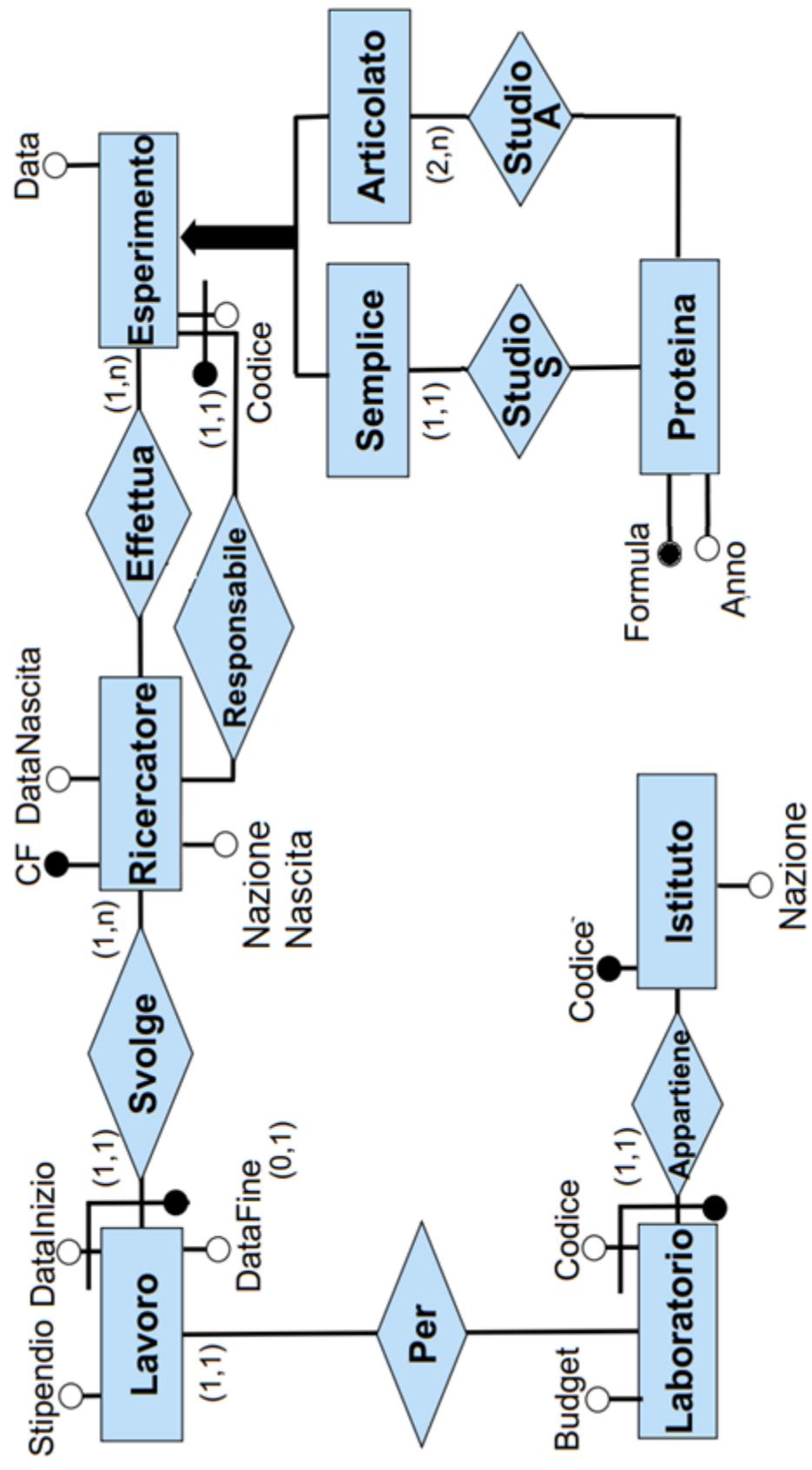
Si noti che un ricercatore non può lavorare contemporaneamente in più di un laboratorio e che non ci sono limiti al numero di rapporti che un ricercatore può avere con i singoli laboratori. Di ogni laboratorio interessa l'istituto a cui appartiene, il codice (unico nell'ambito dell'istituto a cui appartiene) ed il budget attuale. Di ogni istituto interessa il codice (identificativo) e la nazione in cui si trova.

Di ogni esperimento interessano i ricercatori che lo hanno effettuato, tra questi quello che ne è responsabile, il codice (unico per il ricercatore responsabile) e la data in cui si è svolto.

Esistono esattamente due tipi di esperimenti: semplici ed articolati. Ogni esperimento semplice studia singole proteine. Ogni esperimento articolato studia l'interazione tra almeno due proteine.

Ogni proteina è caratterizzata dall'anno in cui è stata scoperta, e dalla sua formula chimica, univoca per una data proteina.

Soluzione Esercizio 1



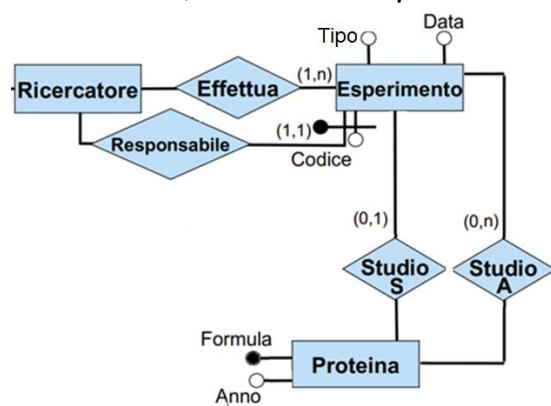
Cardinalità (0,n) sono omesse

Esercizio 2: Progettazione Logica (7 punti)

A partire dallo ER dell'Esercizio 1, produrre uno schema relazionale del database nel riquadro sottostante che minimizza il numero di tabelle. Indicare i vincoli e gli attributi che ammettono valori nulli. Illustrare come ristrutturare il diagramma ER per essere direttamente traducibile in uno schema relazionale.¹

Soluzione Esercizio 2

La fase della ristrutturazione si focalizza sulla eliminazione dei costrutti ER non esprimibili in uno schema relazionale, come attributi multi-valore e generalizzazioni. Nello specifico, occorre eliminare una generalizzazione. Siccome l'obiettivo è minimizzare il numero delle tabelle, le entità *Semplice* e *Articolato* sono accorpate nel padre, con le relazioni collegate. Per differenziare le instanze di un esperimento semplice da quelle di esperimenti articolati, si aggiunge attributo *Tipo*, che può idealmente assumere due valori: *Semplice* e *Articolato*.



Nell'accorpamento, le relazioni *StudioS* e *StudioA* hanno cardinalità minima uguale a 0, perché, se l'esperimento è semplice, si popolerà la prima, altrimenti, se è articolato, si popolerà la seconda. La figura a lato mostra la ristrutturazione del diagramma ER per la parte rilevante, con alcuni attributi omessi per leggibilità. Sebbene non in linea con la metodologia, è anche ammesso accorrere *StudioS* e *StudioA* in una sola relazione *Studio* con cardinalità (1, n) dal lato di *Esperimento*.

Lo schema logico è quindi come segue, dove l'asterisco segue gli attributi che ammettono valori nulli:

PROTEINA(*Formula*, *Anno*)
RICERCATORE(*CF*, *DataNascita*, *NazioneNascita*)
ESPERIMENTO(*Codice*, *Responsabile*, *Data*, *Tipo*, *Proteina**)

Esperimento.Responsabile → *Ricercatore.CF*

Esperimento.Proteina → *Proteina.Formula*

EFFETTUA(*Ricercatore*, *CodiceEsp*, *ResponsabEsp*)

Effettua.Ricercatore → *Ricercatore.CF*

Effettua.(CodiceEsp, ResponsabEsp) → *Esperimento.(Codice, Responsabile)*

STUDIOA(*Proteina*, *CodiceEsp*, *ResponsabEsp*)

Effettua.(CodiceEsp, ResponsabEsp) → *Esperimento.(Codice, Responsabile)*

Effettua.Proteina → *Proteina.Formula*

ISTITUTO(*Codice*, *Nazione*)

LABORATORIO(*Codice*, *Istituto*, *Budget*)

Laboratorio.Istituto → *Istituto.Codice*

LAVORO(*Ricercatore*, *DataInizio*, *LabCodice*, *LabIstituto*, *Stipendio*, *DataFine**)

Lavoro.Ricercatore → *Ricercatore.CF*

Lavoro.(LabCodice, LabIstituto) → *Laboratorio.(Codice, Istituto)*

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 con le relazioni:

- **Diplomato(codice, voto, codScuola, anno)**, che per ogni studente diplomato memorizza il codice, il voto conseguito alla maturità, l'anno di conseguimento e la scuola in cui l'ha conseguita.
- **Scuola(codScuola, citta)** che memorizza codice e città delle scuole.

A. Restituire le città in cui nessuna scuola ha avuto diplomati con 100 nel 2021 (2 punti).²

CITTA_CON_100_NEL_2021=

$\Pi_{\text{citta}} (\sigma_{\text{Voto}=100 \text{ AND } \text{anno}=2021} (\text{Diplomato} \bowtie \text{Scuola}))$

$\Pi_{\text{citta}} (\text{SCUOLA}) - \text{CITTA_CON_100_NEL_2021}$

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

Nome: _____ Num. Matric.: _____

B. Scrivere una query in Standard SQL che, per ogni città con almeno 100 diplomati nel 2021, restituisce il voto medio nel 2021. (2.5 punti).

```
SELECT CITTA, AVG(VOTO)
FROM DIPLOMATO D, SCUOLA S
WHERE D.CODSCUOLA=S.CODSCUOLA
      AND ANNO=2021
GROUP BY CITTA
HAVING COUNT(*)>=100
```

C. Nel riquadro, scrivere una query in Standard SQL che restituisce la città con la scuola con più diplomati con 100 (2.5 punti)

```
CREATE VIEW PIU(CODSCUOLA,NUM-100) AS
SELECT CODSCUOLA, COUNT(*)
FROM DIPLOMATO D
WHERE VOTO=100
GROUP BY CODSCUOLA;

SELECT CITTA
FROM PIU P JOIN SCUOLA S ON
P.CODSCUOLA=S.CODSCUOLA
WHERE P.NUM-100 =
      (SELECT MAX(NUM-100) FROM PIU)
```

Esercizio 4: Transazioni (5 punti)

Si consideri il seguente schedule:

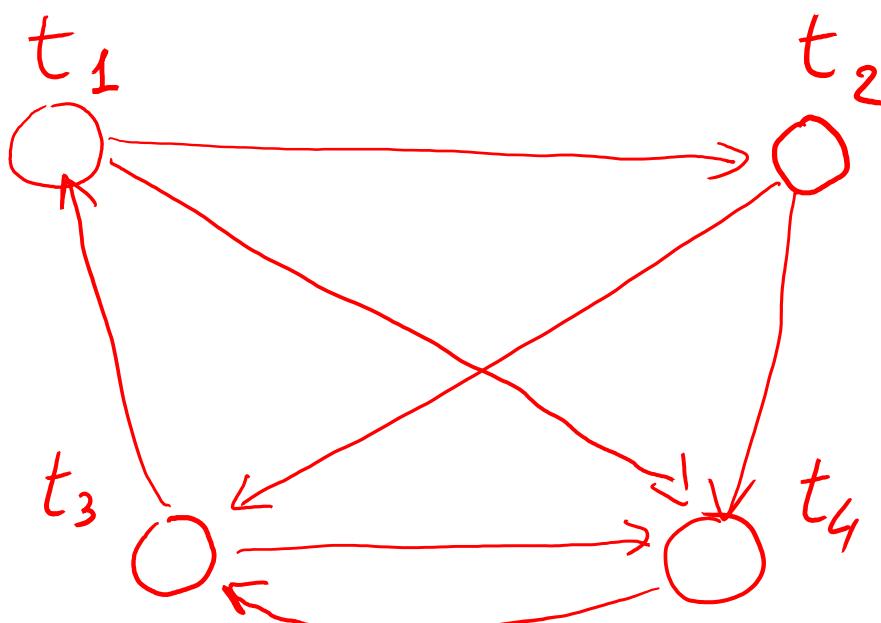
$$S = r2(v) \ r1(x) \ w4(v) \ w3(y) \ w2(x) \ r1(y) \ w4(y) \ w3(v).$$

Dire se S è conflict-serializable o no.

Se la risposta è sì, mostrare uno schedule seriale conflict-equivalente a S.

Se la risposta è no, dire quale transazione causa il problema e come è possibile risolvere il problema.

Il grafo dei conflitti è come seguente:



È chiaramente possibile identificare diversi cicli (t_1 , t_2 e t_3) oppure (t_3 e t_4). Quindi, lo schedule non è conflict-serializzabile.

La transazione t_3 è la transazione che causa i problemi. Il problema può essere risolto posticipando la scrittura di y da parte di t_3 dopo tutte le operazioni delle altre transazioni, ma comunque prima di $w3(v)$:

$$r2(v) \ r1(x) \ w4(v) \ w2(x) \ r1(y) \ w4(y) \ w3(y) \ w3(v)$$

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 A=4 ORDER BY B` sulla relazione $S(A, B, C)$. Quale dei seguenti indici in genere assicura le migliori performance in termini di velocità dell'esecuzione della query?

1. Indice B+Tree sulla coppia (B,A)
2. Indice B+Tree su B
3. Indice B+Tree su A
4. Indice B+Tree sulla coppia (A,B)

Domanda 2 (1.5 Punti)

Sia a data la relazione $R(A,B,C,D)$ con dipendenze funzionali $\{A \rightarrow B, AB \rightarrow C, D \rightarrow AC\}$. Indicare se AD è chiave o superchiave.

1. AD è chiave e superchiave
2. AD è chiave ma non è superchiave
3. AD non è chiave e ma è superchiave
4. AD non è chiave, né superchiave.

Domanda 3 (2 Punti)

Data una relazione $R(\underline{A},B,C,D)$ con dipendenze funzionali $\{A \rightarrow B, AB \rightarrow C, D \rightarrow AC\}$

Quale delle seguenti affermazioni è vera?

1. R è in BCNF **ma non** in 3NF
2. R **non** è in BCNF **ma è** in 3NF
3. R è in BCNF **ed è** in 3NF
4. R non è in BCNF e non è 3NF

Nome: _____ Num. Matric.: _____



Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 5 Settembre 2022

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

Esercizio 1: Diagramma ER (8 punti)

Si richiede di produrre lo schema concettuale Entità-Relazione di un'applicazione relativa agli ospedali ed ai ricoveri. Disegnare il diagramma ER nel riquadro della pagina che segue.

Di ogni ospedale interessa il codice (identificativo), il numero di dipendenti (ma non sempre questa informazione è disponibile), la regione in cui si trova ed i reparti di cui è composto. Di ogni reparto interessa il codice (unico nell'ambito dell'ospedale in cui si trova), la dimensione in metri quadrati.

Le persone vengono ricoverate nei reparti. Di ogni persona interessa il codice fiscale (identificativo), il nome, il cognome e la data di nascita. Per ogni ricovero, si vuole conoscere la persona che viene ricoverata, il reparto in cui viene ricoverato ed il posto occupato.

Si noti che una persona può essere ricoverata più volte in un reparto, ma in date diverse, e non può essere ovviamente ricoverata allo stesso momento in più reparti.

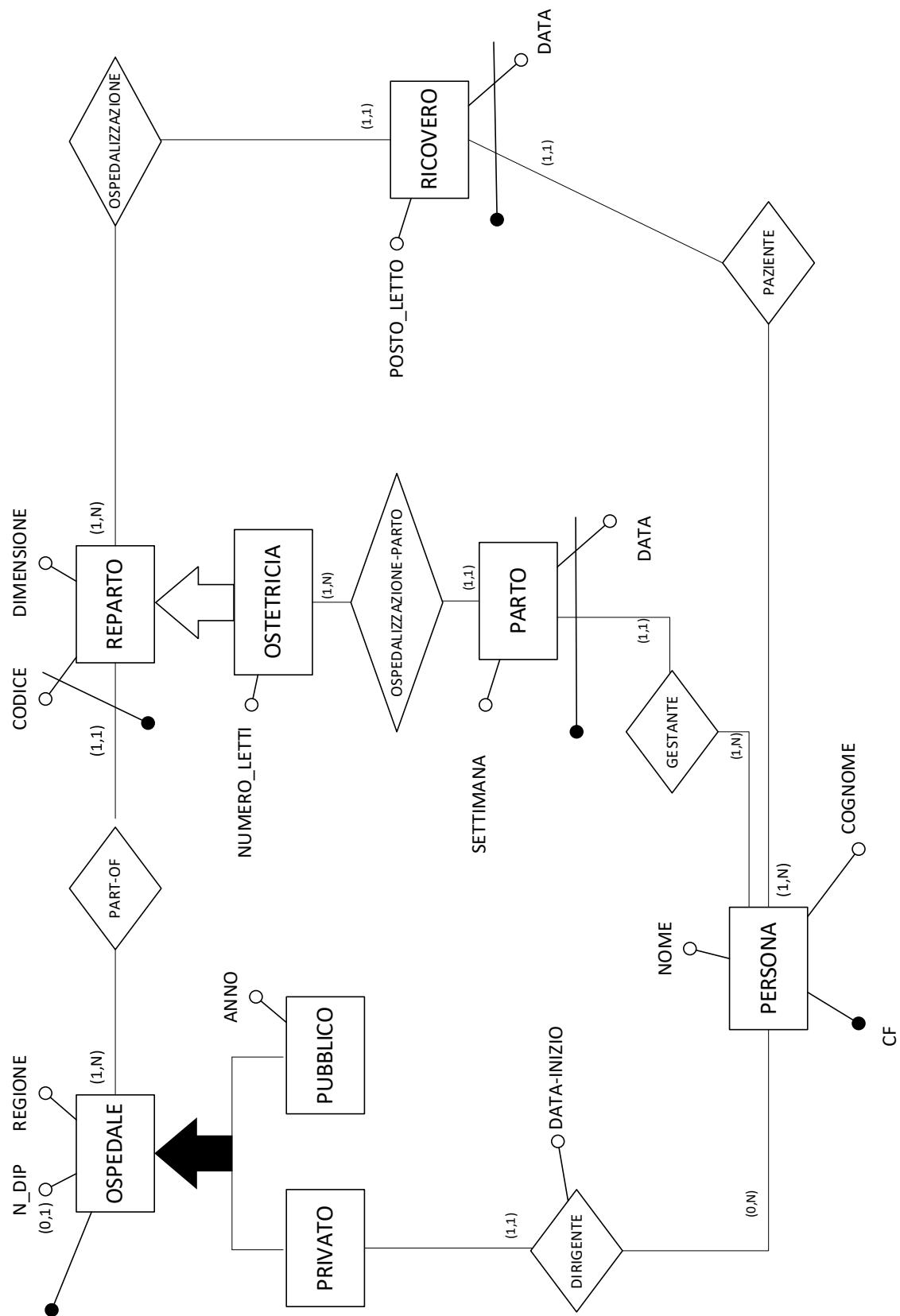
Ci sono ospedali pubblici e privati:

- Dei primi interessa anche l'anno di fondazione
- Dei secondi, in aggiunta alle informazioni di ogni ospedale, interessare sapere chi è la persona dirigente attuale, con la data di inizio mandato.

Gli ospedali hanno anche reparti di ostetricia, per il quali si vuole anche conoscere il numero di posti letto e i parti. Per i parti, si vogliono conoscere la data, la persona gestante, il reparto di ostetricia in cui avviene e la settimana di gestazione in cui il parto accade. La stessa gestante può dare alla luce più neonati nello stesso reparto di ostetricia, ma in date diverse.

Nome: _____ Num. Matric.: _____

Soluzione Esercizio 1

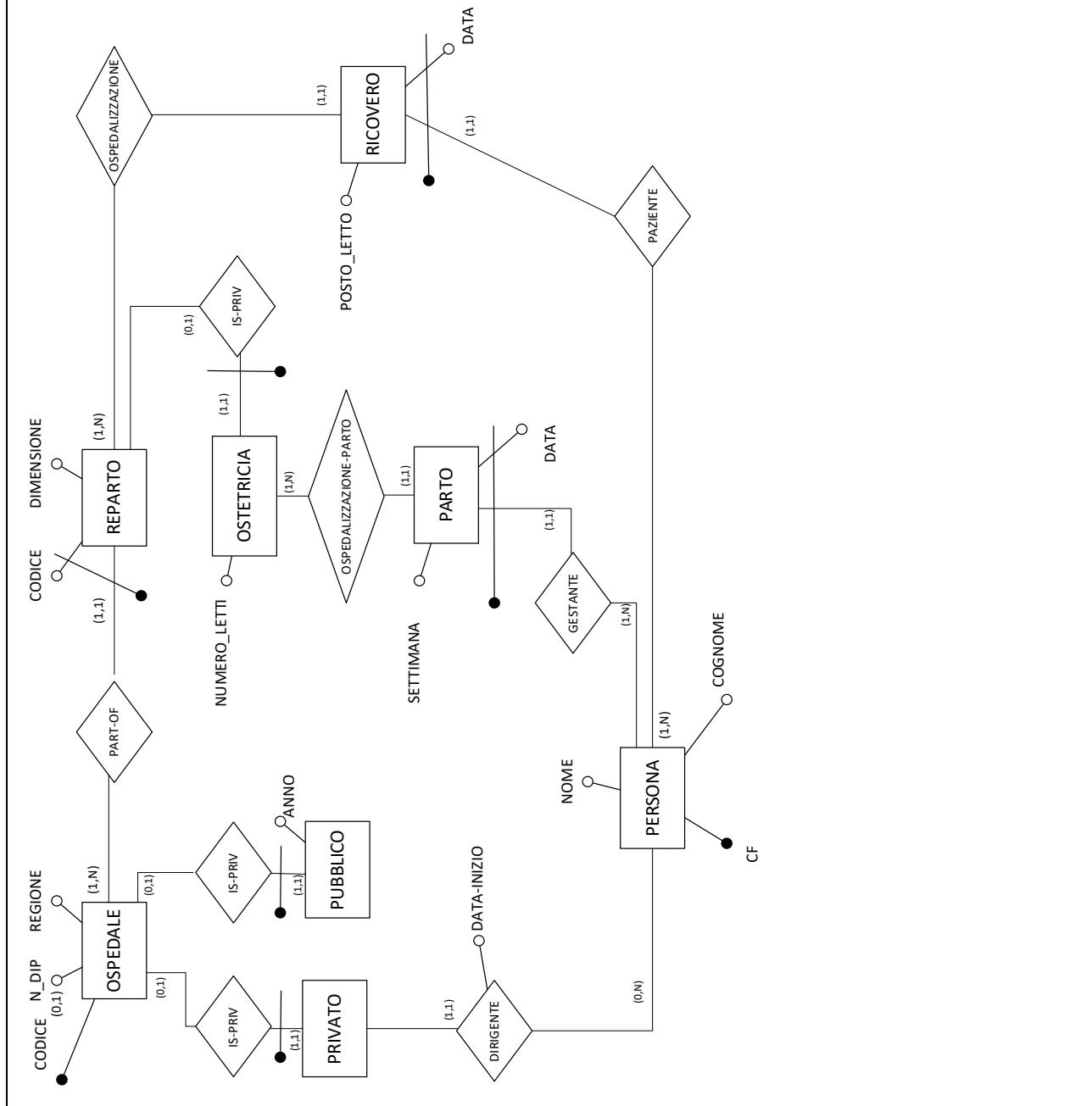


Esercizio 2: Progettazione Logica (7 punti)

A partire dallo ER dell'Esercizio 1, produrre uno schema relazionale del database nel riquadro sottostante che minimizza i valori nulli. Indicare i vincoli e gli attributi che ammettono valori nulli. Illustrare anche come ristrutturare il diagramma ER per essere direttamente traducibile in uno schema relazionale.¹

Soluzione Esercizio 2 (parziale)

È possibile minimizzare la presenza di valori nulli nella ristrutturazione, sostituendo le generalizzazioni con opportune relazioni, come mostrato sotto:



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!).

Nome: _____ Num. Matric.: _____

Esercizio 3: Algebra Relazionale & SQL (7 punti)

Si consideri la seguente base di dati con le relazioni:

- **Fantini** (Nome, Peso, DataNascita)
- **Cavalli** (Nome, AnnoNascita, Scuderia)
- **Gare** (Codice, Nome, Luogo, Data, CavalloVincente, FantinoVincente)

dove $\text{Gare.CavalloVincente} \rightarrowtail \text{Cavalli.Nome}$,

$\text{Gare.FantinoVincente} \rightarrowtail \text{Fantini.Nome}$

A. Il nome dei fantini che hanno vinto solamente gare con cavalli della stessa scuderia (2 punti).²

$J1 = \text{Gare} \bowtie_{\text{CavalloVincente}=\text{Cavalli.Nome}} \text{Cavalli}$

$J2 = J1$

$\Pi_{\text{FantinoVincente}}(J1) -$

$\Pi_{J1.\text{FantinoVincente}}(J1 \bowtie_{J1.\text{Scuderia} <> J2.\text{Scuderia} \text{ AND } J1.\text{FantinoVincente} = J2.\text{FantinoVincente}} J2)$

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.

B. Per ogni fantino che abbia vinto almeno una gara, fornire il nome del fantino ed anche il codice e nome della prima gara che ha vinto (2.5 punti).³

```
CREATE VIEW PRIMAGARA(FANTINO,DATA) AS  
SELECT FANTINOVINCENTE, MIN(DATA)  
FROM GARE  
GROUP BY FANTINOVINCENTE;  
  
SELECT FANTINO, P.DATA, G.CODICE  
FROM PRIMAGARA AS P, GARE AS G  
WHERE P.DATA=G.DATA  
  
AND FANTINOVINCENTE=FANTINO
```

C. Restituire l'anno di nascita medio dei cavalli con cui ha vinto il fantino più giovane – cioè con la data di nascita più grande (2.5 punti).³

```
SELECT AVG(C.ANONASCITA)  
FROM FANTINI F, CAVALLI C, GARE G  
WHERE F.NOME=G.FANTINOVINCENTE AND  
C.NOME=G.CAVALLOVINCENTE AND  
F.DATANASCITA =  
(SELECT MAX(DATANASCITA) FROM FANTINI)
```

³ Se rilevante, si assuma che non ci siano due gare nella stessa data o due fantini nati nello stesso giorno.

Nome: _____ Num. Matric.: _____

Esercizio 4: Decomposizione in Terza Forma Normale (5 punti)

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

- a. Trovare la/e chiave/i di R , motivando la risposta.
- b. Indicare quali dipendenze violano la 3NF, motivando la risposta
- c. Effettuare una decomposizione in 3NF, indicando le chiavi delle relazioni ottenute
- d. La decomposizione preserva tutte le dipendenze? Motivare la risposta.

Punto a

Si considerano le seguenti chiusure

$$\begin{aligned} C^+ &= \{C, A, D, B, E, F\} & A^+ &= \{A, B, F, D, E\} \\ D^+ &= B^+ = F^+ = \{D, B, F, E\} \end{aligned}$$

Quindi, C è una chiave della relazione

Punto b

$A \rightarrow B$ viola perché A non è superchiave, né B è parte di una chiave.

$D \rightarrow B$ e $D \rightarrow E$ violano perché D non è superchiave, e né B né E sono parte di una chiave

$B \rightarrow F$ viola perché B non è superchiave e F non è parte di una chiave.

Punto c

La dipendenza funzionale $C \rightarrow D$ è ridonante poiché $C \rightarrow A, A \rightarrow B, B \rightarrow F, F \rightarrow D$ ed è rimossa dalla copertura ridotta

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

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

2. Viene costruita una relazione per ogni sottoinsieme:

$$R1(C,A), R2(D,B,E,F), R3(A,B)$$

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

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 accadono le condizioni 3 e 4

5. Indicare le chiavi delle relazioni ottenute dalla normalizzazione

$R1(C,A)$ con chiave $\{C\}$, $R2(D,B,E,F)$ con due chiavi: $\{B\}$ o $\{F\}$, $\{D\}$ e $R3(A,B)$ con chiave $\{A\}$. Per $R2$, la soluzione con due chiavi garantisce 3NF e va bene come soluzione: aggiungendo sia $\{B\}$ o $\{F\}$ si ottiene BCNF.

Punto d

No, perché $C \rightarrow D$ non è preservata, poiché non c'è nessuna relazione che contiene entrambi gli attributi. Nel caso errato in cui $C \rightarrow D$ non è rimossa dalla copertura ridotta, $R1$ è $R1(C,A,D)$, che preserva anche $C \rightarrow D$. Tuttavia, la decomposizione non è corretta perché non minimale in termini di attributi

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 MIN(A) FROM S WHERE B=4` sulla relazione $S(A, B, C)$. Quale dei seguenti indici in genere assicura le migliori performance in termini di velocità dell'esecuzione della query?

1. Indice B+Tree sulla coppia (B,A)
2. Indice Hash su B
3. Indice B+Tree su B
4. Indice B+Tree sulla coppia (A,B)

Domanda 2 (1.5 Punti)

Si consideri il seguente log:

CK(T2), B(T5), B(T6), U(T5,O5,B5,A5), C(T5), B(T7), U(T7,O6,B6,A6),
B(T8), U(T6,O1,B7,A7), A(T7), C(T6), guasto

Di quali transazioni occorre fare lo UNDO?

1. T2,T5,T7,T8
2. T2,T5,T7
3. T2,T7,T8
4. T5,T6

Domanda 3 (2 Punti)

Data una relazione $R(A,B,C,D)$ con dipendenze funzionali $\{A \rightarrow BC, C \rightarrow AD\}$

Quale delle seguenti affermazioni è vera?

1. R è in BCNF **ma non** in 3NF
2. R **non** è in BCNF **ma è** in 3NF
3. R è in BCNF ed è in 3NF
4. R **non** è in BCNF **e non** è 3NF

Nome: _____ Num. Matric.: _____



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 21 Settembre 2022

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**

Esercizio 1: Diagramma ER (8 punti)

Si richiede di produrre lo schema concettuale Entità-Relazione di un'applicazione relativa agli interventi di manutenzione su aerei. Disegnare il diagramma ER nel riquadro della pagina che segue.

Di ogni aereo interessa il codice (identificatore), il modello, la ditta che l'ha costruito, e l'anno di costruzione. Gli aerei da turismo sono particolari aerei di cui interessa anche il costo. Di ogni ditta interessa il codice fiscale (identificatore), la città in cui ha sede

Gli aerei sono soggetti ad interventi di manutenzione: di ogni intervento interessa l'aereo oggetto dell'intervento, la data, il costo di manodopera dell'intervento e la città in cui si è svolto. Un aereo non può essere oggetto di più di un intervento di manutenzione al giorno. Si noti che il costo della manodopera non è sempre noto dall'inizio: viene tipicamente calcolato alla fine dell'intervento.

Di ogni intervento di manutenzione che riguarda un aereo da turismo interessa anche la ditta responsabile dell'intervento stesso. Un intervento di manutenzione può richiedere la sostituzione di pezzi di un certo tipo. Di ogni tipo di pezzo, interessano il codice (identificatore), il costo unitario, e la ditta costruttrice. Se il tipo di pezzo è un tipo di pezzo del motore, interessa sapere anche l'indice di affidabilità (intero positivo). Infine, per ogni pezzo oggetto di sostituzione nell'ambito di un intervento di manutenzione interessa sapere la quantità sostituita (ad esempio, in un intervento di manutenzione vengono sostituiti 3 poggiastesta dell'aereo).

Per i pezzi motore di un dato intervento, si vuole sapere lo sconto praticato sul prezzo unitario nell'ambito di quell'intervento.

Nome: _____ Num. Matric.: _____

Soluzione Esercizio 1

Soluzione non fornita

Esercizio 2: Progettazione Logica (7 punti)

A partire dallo ER dell'Esercizio 1, produrre uno schema relazionale del database nel riquadro sottostante che minimizza i valori nulli. Indicare i vincoli e gli attributi che ammettono valori nulli. Illustrare anche come ristrutturare il diagramma ER per essere direttamente traducibile in uno schema relazionale.¹

Soluzione Esercizio 2

Soluzione non fornita

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!)

Nome: _____ Num. Matric.: _____

Esercizio 3: Algebra Relazionale & SQL (7 punti)

Si consideri il sistema per la registrazione degli iscritti a diversi concerti, con una relazione N-a-N tra concerti e persone:

CONCERTO (ID, CANTANTE, CITTA)

PERSONA (PCF, NOMINATIVO, DATA-NASCITA)

SPETTATORE (ID-CONCERTO, CF)

dove CANTANTE è il codice fiscale, con chiavi esterne:

SPETTATORE.CF --> PERSONA.PCF

SPETTATORE.ID-CONCERTO --> CONCERTO.ID

CONCERTO.CANTANTE --> PERSONA.PCF

(viene memorizzato nella relazione concerto il CF del cantante)

Si noti che un cantante può anche partecipare come spettatore ad un concerto.

A. Definire una query in Algebra Relazionale che restituisce i nominativi degli spettatori che hanno partecipato a concerti di almeno due cantanti (2 punti).²

C1 = Concerto $\bowtie_{ID-CONCERTO=ID}$ Spettatore

C2 = C1

$\Pi_{Nominativo} \text{ (Persona}$

$\bowtie_{C1.CF=PCF} (C1 \bowtie_{C1.CF=C2.CF \text{ AND } C1.Cantante \leftrightarrow C2.Cantante} C2)$

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

B. Scrivere una query in standard SQL che restituisce la città in cui è avvenuto il concerto con il numero più alto di spettatori (2.5 punti).

```
CREATE VIEW SPETTATORI-PER-CONCERTO
    SELECT ID-CONCERTO, COUNT(CF) AS NUM
    FROM SPETTATORE
    GROUP BY ID-CONCERTO

    SELECT CITTA, NUM
    FROM SPETTATORI-PER-CONCERTO S, CONCERTO
    WHERE S.ID-CONCERTO = CONCERTO.ID AND
        NUM=
            (SELECT MAX(NUM)
            FROM SPETTATORI-PER-CONCERTO)
```

C. Scrive una query in standard SQL che, per ogni cantante, restituisce il CF del cantante e il numero medio di spettatori ai suoi concerti (2.5 punti).

Si usa la stessa view SPETTATORI-PER-CONCERTO definita sopra.

```
SELECT CANTANTE, AVG(NUM)
FROM CONCERTO, SPETTATORI-PER-CONCERTO
WHERE ID-CONCERTO=ID
GROUP BY CANTANTE
```

Nome: _____ Num. Matric.: _____

Esercizio 4: Decomposizione in Terza Forma Normale (5 punti)

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

- a. Trovare la/e chiave/i di R , motivando la risposta.
- b. Calcolare la copertura ridotta
- c. Effettuare una decomposizione in 3NF, indicando le chiavi delle relazioni ottenute
- d. Indicare se c'è o ci può essere perdita nel join, motivando la risposta.

Punto a

Si considerano le seguenti chiusure

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

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

Quindi, AB e AC sono chiavi della relazione

Punto b

F viene trasformata in modo tale da avere un solo attributo a destra: $AB\rightarrow C$, $AB\rightarrow D$, $AB\rightarrow E$, $AC\rightarrow B$, $AC\rightarrow D$, $AC\rightarrow E$, $B\rightarrow C$, $C\rightarrow B$, $C\rightarrow D$, $B\rightarrow E$. Siccome c'è $B\rightarrow C$, $AB\rightarrow C$ è ridondante e può essere rimossa dalla copertura ridotta. Lo stesso vale per $AB\rightarrow D$, $AB\rightarrow E$, $AC\rightarrow B$, $AC\rightarrow D$, $AC\rightarrow E$ e. La copertura ridotta risulta quindi essere $\{B\rightarrow C, C\rightarrow B, C\rightarrow D, B\rightarrow E\}$.

Punto c

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

Siccome la chiusura di B e C sono uguali. Si crea un solo gruppo: $\{B\rightarrow C, C\rightarrow B, C\rightarrow D, B\rightarrow E\}$

2. Viene costruita una relazione per ogni sottoinsieme:

$R_1(B, C, D, E)$

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 $T(K)$:

Si aggiungono $R_2(A, B)$ e $R_3(A, C)$.

5. Indicare le chiavi delle relazioni ottenute dalla normalizzazione

R_1 ha due chiavi: $\{B\}$ e $\{C\}$. R_2 ha chiave $\{A, B\}$ e R_3 ha chiave $\{A, C\}$

Punto d

L'attributo comune a R_1 e R_2 è B , che è chiave di R_1 . L'attributo comune a R_1 e R_3 è C , che è chiave di R_3 . L'attributo comune a R_2 e R_3 è A , che non è chiave né di R_2 né di R_3 . Quindi, ci può essere perdita nel join.

Esercizio 5: Quiz (5 punti)

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

Domanda 1 (21.5 Punti)

Data la relazione R(A, B, C, D) con dipendenze funzionali $AB \rightarrow C$ e $C \rightarrow D$, indicato con $|R|$ il numero di tuple di $R(A, B, C, D)$. Quale affermazione è vera per l'operazione $\pi_{A,B}(R)$ in Algebra Relazionale?

1. $\pi_{A,B}(R) = |R|$
2. $\pi_{A,B}(R) < |R|$
3. $\pi_{A,B}(R) \leq |R|$ e, in certa istanze valide di R , può accadere che $\pi_{A,B}(R) < |R|$
4. $\pi_{A,B}(R) = 0$

Domanda 2 (1.5 Punti)

Si consideri il seguente log:

CK(T2), B(T5), B(T6), U(T5,O5,B5,A5), C(T5), B(T7), U(T7,O6,B6,A6),
B(T8), U(T6,O1,B7,A7), A(T7), C(T6), guasto

Sapendo che occorre fare l'UNDO di T2, T7 e T8, e il REDO di T5 e T6. Quale è la prima operazione da effettuare per la ripresa a caldo?

1. O5=A5
2. O1=A7
3. O1=B7
4. O6=B6

Domanda 3 (32 Punti)

Si consideri lo schedule $S = r1(a), r2(b), w1(b), r2(a) w2(a) r3(a) w3(a)$

Indicare se S è view-serializzabile (VSR) e/o conflict-serializzabile (CSR).

1. R è in VSR **ma non** in CSR
2. R **non** è in VSR **ma è** in CSR
3. R è in VSR **ed è** in CSR
4. R **non** è in VSR **e non è** CSR

Nome: _____ Num. Matric.: _____



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 6 Febbraio 2023

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**

Esercizio 1: Diagramma ER (8 punti)

Si richiede di produrre lo schema concettuale Entità-Relazione di un'applicazione relativa per un'applicazione relativa ad un negozio di noleggio motociclette. Disegnare il diagramma ER nel riquadro della pagina che segue.

Il negozio possiede tante motociclette. Di ogni motocicletta, interessa la targa(identificativo), il modello, l'anno di immatricolazione ed i chilometri percorsi. Di ogni modello interessala marca (ad esempio, "Honda"), il codice (unico nell'ambito della marca), la cilindrata, il prezzo del noleggio all'ora ed il prezzo del noleggio al chilometro.

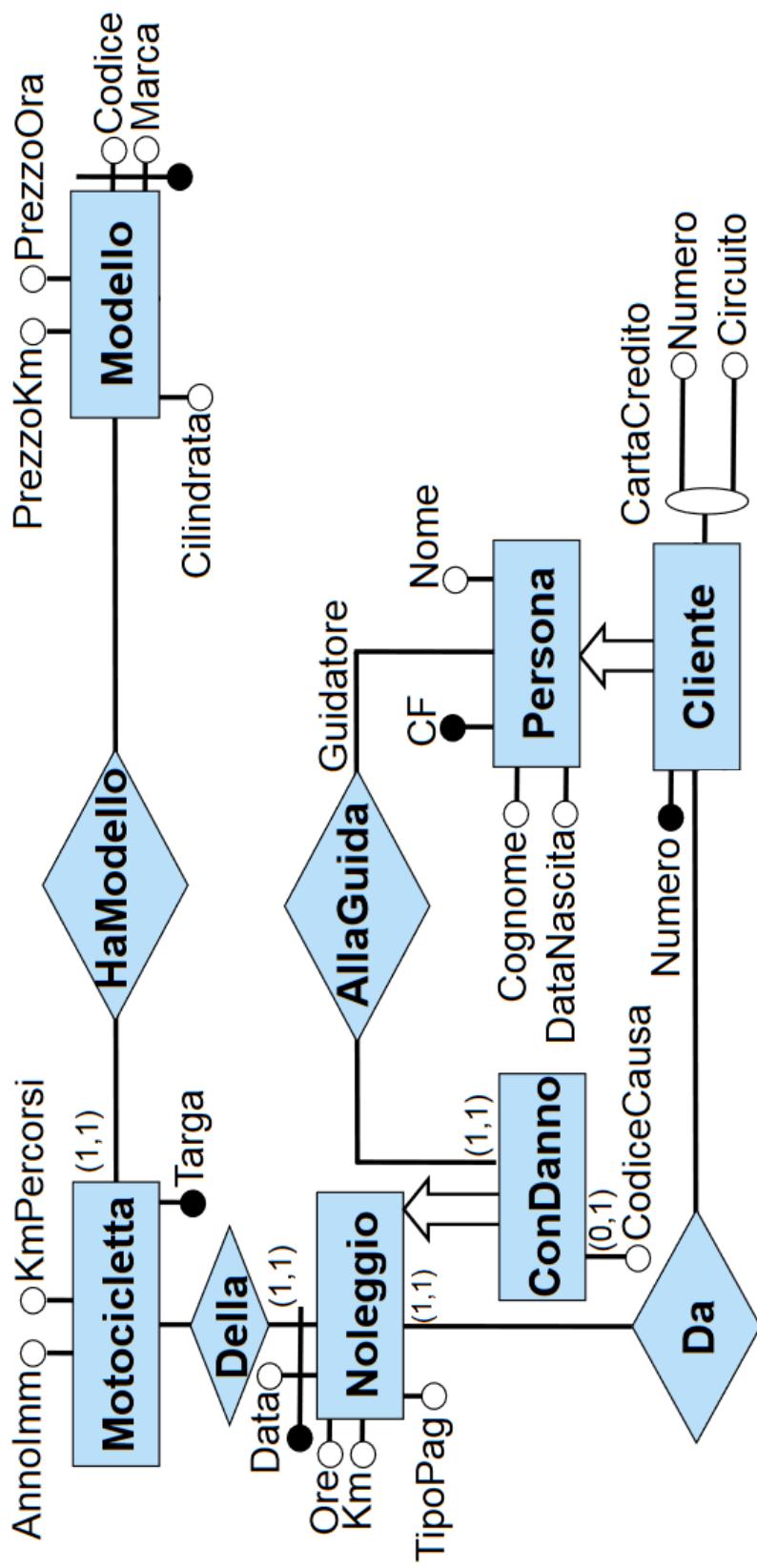
Di ogni noleggio, interessa il cliente che l'ha effettuato, la motocicletta noleggiata, la data in cui è avvenuto (ogni noleggio è infatti relativo ad un giorno), la durata in ore, la tipologia di pagamento (a chilometro oppure ad ora) ed i chilometri percorsi. Si noti che una stessa motocicletta non può essere noleggiata più volte nello stesso giorno.

Dei clienti, Il negozio vuole conoscere il numero di telefono. Ogni cliente possiede una carta di credito, di cui interessa il numero ed il circuito.

Per quei noleggi per i quali si è verificato un danno alla motocicletta noleggiata, interessa il codice indicante la causa del danno (se nota) e la persona-conducente che era alla guida della motocicletta al momento del verificarsi del danno. Di ogni conducente, l'applicazione deve memorizzare i seguenti dati: codice fiscale (identificativo), nome, cognome e data di nascita. Si noti che il conducente può coincidere con il cliente del noleggio oppure essere un'altra persona.

Nome: _____ Num. Matric.: _____

Soluzione Esercizio 1



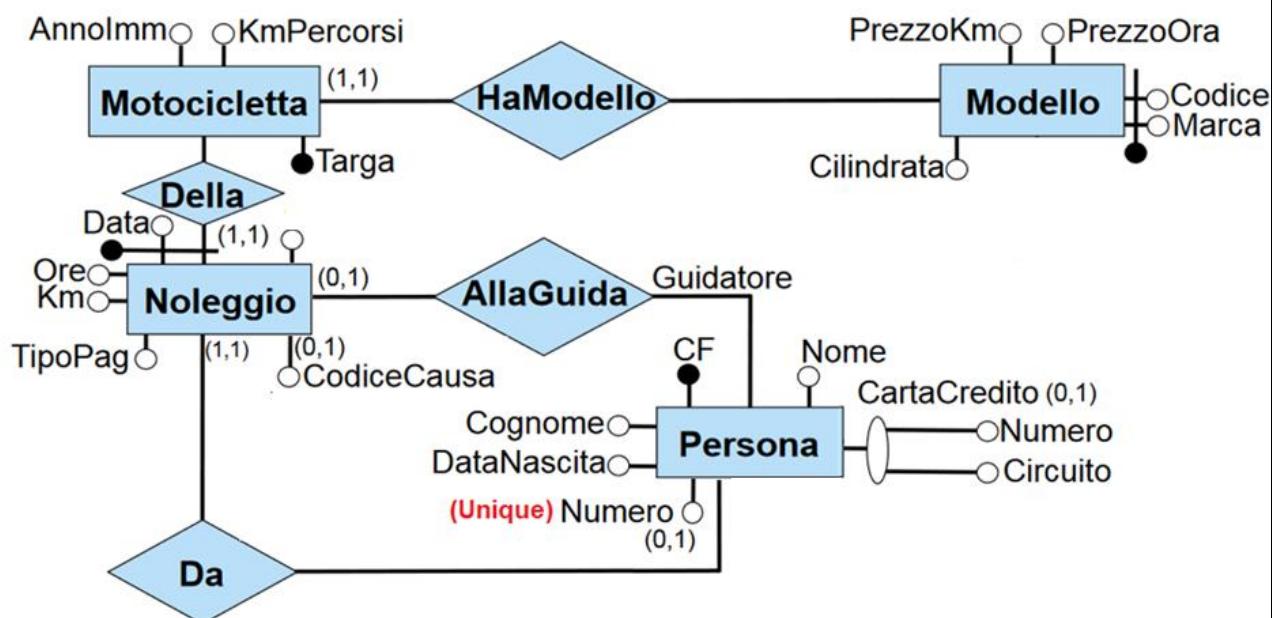
Cardinalità (0,n) sono omesse

Esercizio 2: Progettazione Logica (7 punti)

A partire dallo ER dell'Esercizio 1, produrre uno schema relazionale del database nel riquadro sottostante che minimizza i valori nulli. Indicare i vincoli e gli attributi che ammettono valori nulli. Illustrare come ristrutturare il diagramma ER per essere direttamente traducibile in uno schema relazionale.¹

Soluzione Esercizio 2

Il primo passaggio è la ristrutturazione dello ER per rimuovere le generalizzazioni. Siccome l'obiettivo è minimizzare il numero di tabelle, l'entità *ConDanno* viene fusa con *Noleggio*, l'entità *Cliente* viene fusa con *Persona*. Questo ovviamente introduce un certo numero di attributi con cardinalità (0,1), cioè che ammettono il valore nullo nelle corrispondenti tabelle. Il diagramma ristrutturato è quindi il seguente:



Il vincolo unique per numero serve a rappresentare che, se una persona ha un numero telefonico, allora è l'unica con quel numero. Questo viene derivato dal fatto che l'entità *Cliente* nel diagramma ER concettuale usa tale attributo come identificatore. Si noti che una persona potrebbe non avere un numero di telefono perché solo le persone clienti lo hanno

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!).

Nome: _____ Num. Matric.: _____

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 (Città, Servizio, Anno, Spesa)**
- **POSIZIONE (Città, 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. Scrivere una query in Algebra Relazione che restituisce le città che forniscono almeno due servizi sociali nel 2021 (2 punti).²

$S2021 = \sigma_{Anno=2021} SERVIZI_SOCIALI$

$T2021 = S2021$

$\pi_{S2021.Città} ($

$S2021 \bowtie_{S2021.Città=T2021.Città \text{ AND}}$
 $S2021.SERVIZIO <> T2021.SERVIZIO T2021)$

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.

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

```
CREATE VIEW SPESA_CITTA_2020(CITTA, SPESA) AS
    SELECT CITTA, SUM(SPESA)
    FROM SERVIZI_SOCIALI
    WHERE ANNO=2020
    GROUP BY CITTA

    SELECT REGIONE
    FROM POSIZIONE AS P , SPESA_CITTA_2020 AS S
    WHERE P.CITTA=S.CITTA AND SPESA=
        SELECT MIN(SPESA) FROM SPESA_CITTA_2020
```

C. Nel riquadro, scrivere una query in Standard SQL che, per ogni regione, restituisce la regione e il numero medio dei servizi offerti dalle città in SERVIZI_SOCIALI nel 2020 (2.5 punti)

```
SELECT REGIONE, AVG(SPESA)
FROM SERVIZI_SOCIALI S, POSIZIONE P
WHERE S.CITTA=P.CITTA AND ANNO=2020
GROUP BY REGIONE
```

Nome: _____ Num. Matric.: _____

Esercizio 4: Decomposizione in Terza Forma Normale (5 punti)

Sia data la seguente relazione R(ABCDE) con copertura ridotta

$$G=\{B \rightarrow C, B \rightarrow E, C \rightarrow A \text{ e } C \rightarrow D\}.$$

Risolvere i seguenti punti, motivando le risposte:

- a. Trovare la/e chiave/i di R, motivando la risposta.
- b. Decomporre in 3NF, indicare le chiavi delle relazioni finali ottenute.
- c. La decomposizione ottenuta al punto b è anche in BCNF rispetto a G?
- d. Indicare se c'è o ci può essere perdita nel join.

a. La chiusura di B è $B^+ = \{A, B, C, D, E\}$. La chiusura di C è $C^+ = \{A, C, D\}$. Quindi, B è superchiave (ed anche chiave), mentre C non è superchiave poiché mancano B ed E in C^+ .

b. Siccome la copertura ridotta è già data come testo dell'esercizio, occorre solamente fare i seguenti passi:

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

Si ottiene un insieme $\{B \rightarrow C, B \rightarrow E\}$ ed un secondo $\{C \rightarrow A, C \rightarrow D\}$.

2. Viene costruita una relazione per ogni sottoinsieme: $R1(B, C, E)$ e $R2(C, B, D)$. Siccome ogni relazione ha una sola chiave, questa viene rappresentata sottolineata.

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

La condizione non si verifica; quindi, rimangono $R1$ e $R2$

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

La condizione non si verifica; quindi, rimangono $R1$ e $R2$

Parte c

Lo schema $R1(B, C, E)$ e $R2(C, A, D)$ è in BCNF:

- $B \rightarrow C, B \rightarrow E$ sono le dipendenze relative a $R1$. La parte di sinistra è B che è superchiave di $R1$ (anche chiave)
- $C \rightarrow A, C \rightarrow D$ sono le dipendenze relative a $R2$. La parte di sinistra è C che è superchiave di $R2$ (anche chiave)

Parte d

Gli attributi in comune tra $R1$ e $R2$ consistono nel solo C. C è chiave di $R2$. Siccome condizione sufficiente per una decomposizione senza perdite è che gli attributi in comune siano superchiave di almeno una delle due relazioni, la decomposizione è in effetti senza perdita.

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 A=4 AND B=7` sulla relazione $S(A, B, C)$. Quale dei seguenti indici in genere assicura le migliori performance in termini di velocità dell'esecuzione della query?

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

Domanda 2 (1.5 Punti)

Data la relazione $R(A, B, C, D)$ con dipendenze funzionali $A \rightarrow CD$ e $B \rightarrow A$.

Una delle seguenti affermazioni è vera.

1. AB è chiave e superchiave
2. AB è chiave ma non è superchiave
3. AB non è chiave e ma è superchiave
4. AB non è chiave, né superchiave.

Domanda 3 (2 Punti)

Si consideri lo schedule $S = \{r1(a) w2(a) r3(a) w1(b) w3(c) r2(d) r3(d) w3(b) w3(e)\}$

Dopo aver calcolato il grafo dei conflitti, indicare se S è conflict-serializzabile (CSR) e/o view-serializzabile (VSR)?

1. R è in CSR **ma non** in VSR
2. R **non** è in CSR **ma è** in VSR
3. R è in CSR ed è in VSR
4. R **non** è in CSR **e non** è VSR

Nome: _____ Num. Matric.: _____



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 24 Giugno 2024

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”
 - verranno segnalati agli uffici competenti dell'Università

Esercizio 1: Diagramma ER (7 punti)

Si richiede di produrre lo schema concettuale Entità-Relazione di un'applicazione relativa ad un insieme di scuole di addestramento di cani. Disegnare il diagramma ER nel riquadro della pagina che segue.

Nel database, per ogni scuola è di interesse memorizzare la città in cui si trova, il codice (unico nell'ambito della città ma non in assoluto), il nome del direttore, e l'anno di fondazione. Di ogni città interessa il nome (identificativo), la regione ed il numero di abitanti.

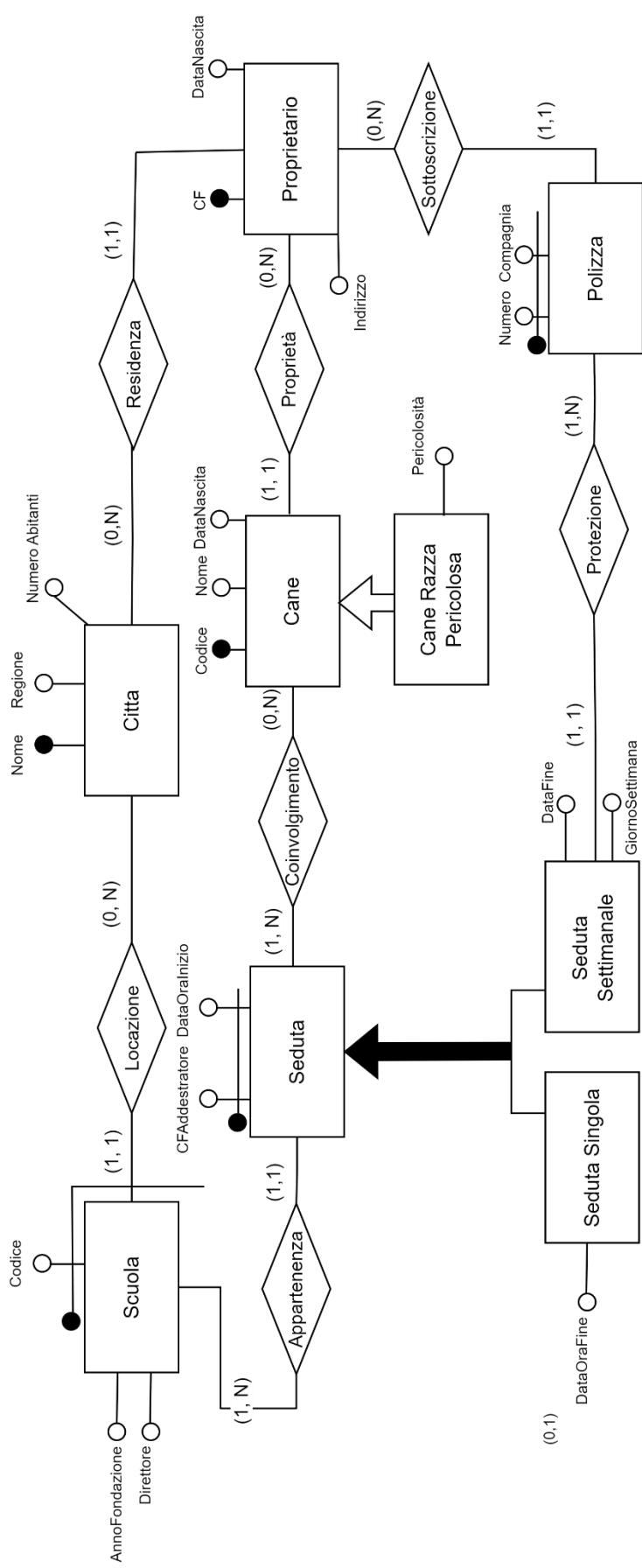
La scuola organizza sedute di addestramento, ognuna identificata dalla data ed ora di inizio e dal codice fiscale dell'addestratore coinvolto. Le sedute di addestramento possono essere singole oppure settimanali: delle prime si memorizza anche la data e ora di fine (che potrebbe non essere nota sin dal primo inserimento), delle seconde la data di fine del ciclo di sedute e il giorno della settimana in cui si ripetono.

Ogni seduta coinvolge uno o più cani. Di ogni cane interessa il codice (identificativo), il nome, la data di nascita ed il proprietario, di cui interessa il codice fiscale – identificativo, la data di nascita, la città e l'indirizzo di residenza. Esistono cani di razze aggressive, di cui interessa anche il livello di pericolosità.

Data la natura non occasionale, ogni seduta settimanale deve essere associata ad una polizza di assicurativa, identificata da un numero e dalla compagnia assicurativa. Ogni polizza è sottoscritta (e quindi associata) ad uno ed uno solo proprietario di cani.

Nome: _____ Num. Matric.: _____

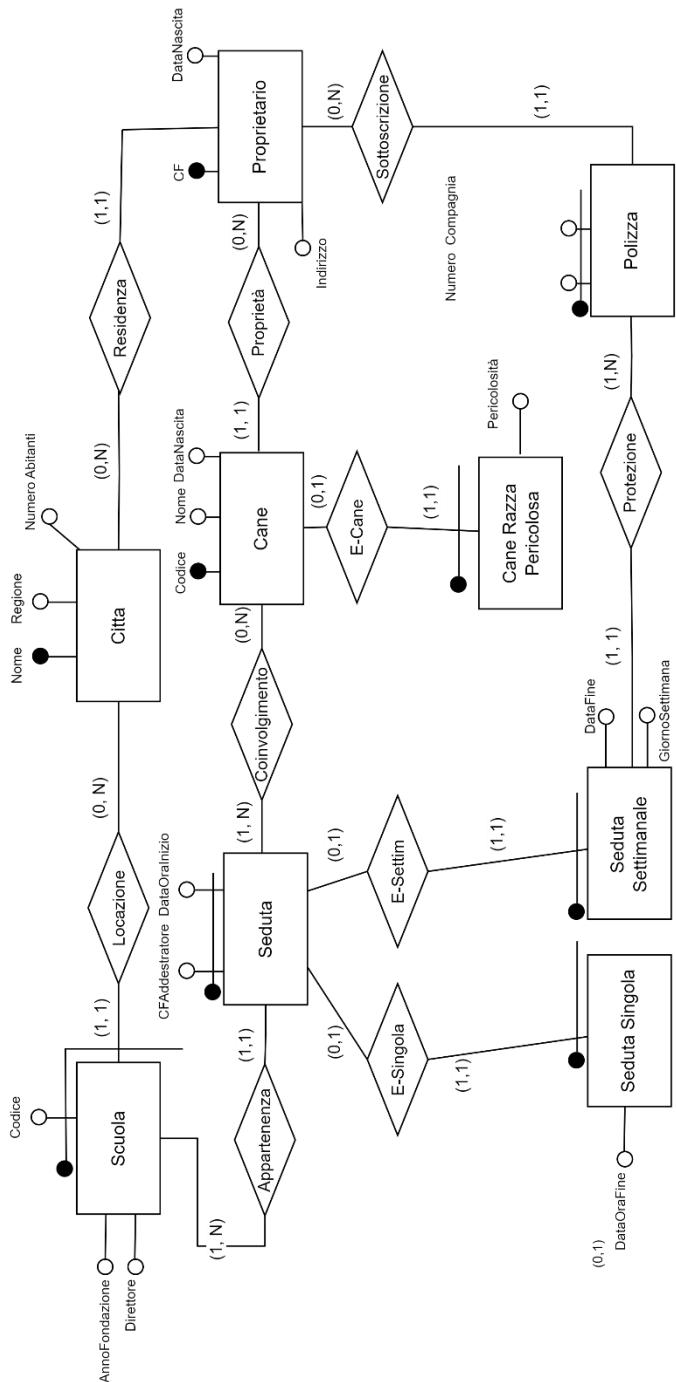
Soluzione Esercizio 1



Esercizio 2: Progettazione Logica (7 punti)

A partire dal Diagramma ER dell'Esercizio 1, produrre uno schema relazionale del database nel riquadro sottostante minimizzando i valori nulli. Indicare i vincoli di chiave e gli attributi che ammettono valori nulli. Illustrare come ristrutturare l'ER per essere traducibile in uno schema relazionale.¹

Soluzione Esercizio 2 (Parziale, solo il diagramma ristrutturato)



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!).

Nome: _____ Num. Matric.: _____

Esercizio 3: Algebra Relazionale & SQL (7 punti)

Si consideri la seguente base di dati per i voli che vengono effettuati nel mondo:

VOLO(IdVolo,GiornoSett,CittàPart,OraPart,CittàArr,OraArr,TipoAereo)

AEREO(TipoAereo,NumPasseggeri,QtaMerci)

A. Restituire il tipo di aereo con il massimo numero di passeggeri, tra quelli in uso per voli che partono da Roma (2 punti).²

$C1 = \sigma_{CittàPart='Roma'}(VOLO \bowtie AEREO)$

$C2 = C1$

$TIPO-AEREO-NON-MASSIMO =$

$\Pi_{C1.TipoAereo} (C1 \bowtie_{C1.NumPasseggeri < C2.NumPasseggeri} C2)$

$(\Pi_{TipoAereo} (C1)) \setminus TIPO-AEREO-NON-MASSIMO$

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. Nel riquadro, scrivere una query in Standard SQL che restituisce il giorno della settimana in cui ci sono più voli in arrivo a Roma. (2.5 punti).

```
CREATE VIEW VOLI-PER-GIORNO-ROMA(GiornoSett,Num) AS
SELECT GiornoSett, COUNT(*)
FROM VOLO WHERE CittàArr='Roma'
GROUP BY GiornoSett

SELECT GiornoSett
FROM VOLI-PER-GIORNO-ROMA
WHERE Num=
  (SELECT MAX(Num) FROM VOLI-PER-GIORNO-ROMA)
```

C. Nel riquadro, scrivere una query in Standard SQL che restituisce, per ogni città C da cui partono almeno 100 voli settimanalmente, restituisce il numero medio di aerei che giornalmente partono da C (2.5 punti).³

```
SELECT CittaPart, Count(*)/7
FROM VOLO
GROUP BY CittaPart
HAVING Count(*)>=100
```

³ Vengono restituite coppie del tipo (C,N) dove C è la città e N è il numero medio di aereo che giornalmente partono da C

Esercizio 4: Normalizzazione (6 punti)

Sia data la relazione $R(A,B,C,D,E,F)$ con copertura ridotta $G=\{C \rightarrow B, C \rightarrow D, C \rightarrow F, BE \rightarrow C, B \rightarrow A, AD \rightarrow E\}$

- a. Trovare la/e chiave/i di R , motivando la risposta.
- b. Quali dipendenze violano la 3NF? Motivare la risposta
- c. Effettuare una decomposizione in 3NF.
- d. La decomposizione è anche in BCNF? Motivare la risposta

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

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 chiavi C e BE .

$R_2(B, A)$ con chiave B

$R_3(AD, E)$ con chiave AD

Punto d

Sì, la decomposizione è anche in BCNF. (Per questioni di spazio, non viene mostrato la spiegazione necessaria, che consiste nel mostrare che ogni relazione non viola BCNF).

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)

Sia data la seguente porzione di log fino al guasto: CK(T5,T6), B(T7), U(T7,O6,B6,A6), U(T6, O3, B7, A7), B(T8), I(T8,O5,A5), C(T8), A(T5). Sapendo che occorre effettuare l'UNDO di T5,T6 e T7 e il REDO di T8, quale è la prima operazione da effettuare per la ripresa a caldo?

1. D(O5)
2. O3=B7
3. O6=B6
4. O5=A5

Domanda 2 (1.5 Punti)

Data la relazione $R(A, B, C, D)$ e $S(W, X, Y, Z)$, indicato con $|R|$ e $|S|$ il numero di tuple di R e S . Sapendo che non ci sono chiavi esterne, quale affermazione è vera per l'operazione $S \bowtie_{A=W} R$ in Algebra Relazionale?

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

Domanda 3 (2 Punti)

Si consideri le relazioni $R(A, B, C, D)$ e la seguente query

SELECT * FROM R WHERE A=10 AND B>7

Quale dei seguenti indici garantisce l'efficienza massima?

1. Indice Hash sulla coppia (B,A)
2. Indice Hash sulla coppia (A,B)
3. Indice B-TREE sulla coppia (A,B)
4. Indice B-TREE sulla coppia (B,A)

Nome: _____ Num. Matric.: _____



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 17 Luglio 2024

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 ripetere il progetto**
 - **verranno segnalati agli uffici competenti dell'Università**

Esercizio 1: Diagramma ER (8 punti)

*Si richiede di produrre lo schema concettuale Entità-Relazione di una base di dati che gestisca i procedimenti sanzionatori ad esercizi commerciali. **Disegnare il diagramma ER nel riquadro della pagina che segue.***

Le autorità di pubblica sicurezza periodicamente effettuano indagini per esercizi commerciali. Ogni indagine è caratterizzata da un nome e da una frequenza di occorrenza.

Come detto, ogni indagine viene periodicamente ripetuta in diverse edizioni. Ogni edizione è identificata dall'indagine di cui è istanza e da un codice univoco tra tutte le edizioni di una stessa indagine (cioè lo stesso codice può essere solamente utilizzato per edizioni di indagini differenti). Per ogni edizione di indagine, si vuole anche memorizzare la data di inizio e di fine dell'edizione dell'indagine.

Per ogni edizione di indagine, le autorità possono sospettare delle irregolarità ed aprire un procedimento. Ogni procedimento si riferisce ad una ed un'una sola edizione di indagine ed è caratterizzato da una data di apertura. Ogni procedimento ha un fascicolo associato ed è di interesse conoscere il soggetto interessato dall'indagine.

Ogni fascicolo è per un solo procedimento ed è univocamente identificato da un codice e dal procedimento a cui è associato. Un fascicolo contiene diversi documenti. Ogni documento è identificato da un codice univoco all'interno del fascicolo. Di ogni documento, è di interesse memorizzare il nome del documento, la data di produzione e, se il documento è disponibile digitalmente, il path da cui scarire il documento stesso.

I soggetti possono essere persone fisiche o imprese. Delle persone fisiche interessa memorizzare il codice fiscale, delle imprese la partita iva. Inoltre, per entrambi, è di interesse l'indirizzo cui il soggetto è contattabile.

Si noti che le imprese possono prevedere delle unità locali, di cui si vuole mantenere informazione nella base di dati: l'impresa può essere composta di altra imprese "periferiche", che a loro volta possono comprendere ulteriori imprese periferiche a cascata.

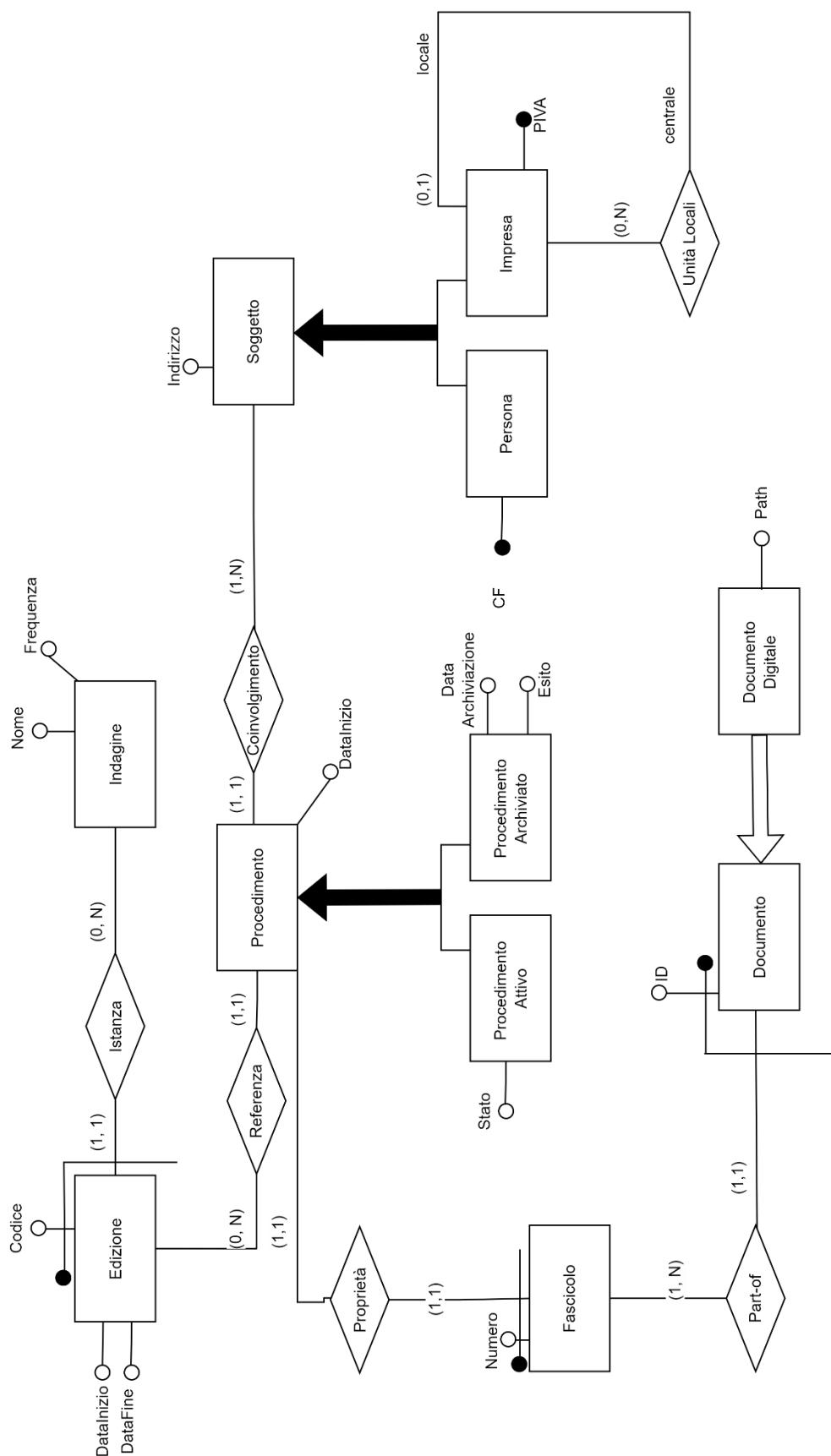
Si vuole anche tenere lo storico dei procedimenti. Un procedimento può essere ancora attivo oppure archiviato a fine indagine. Quando è ancora attivo (cioè non archiviato), si vuole mantenere informazione dello stato; quando i procedimenti vengono archiviati, si vuole sapere l'esito e la data di archiviazione.

Nome: _____

Num. Matric.: _____

Soluzione Esercizio 1

Si noti che il Diagramma ER concettuale non richiede che ogni entità abbia un identificatore.

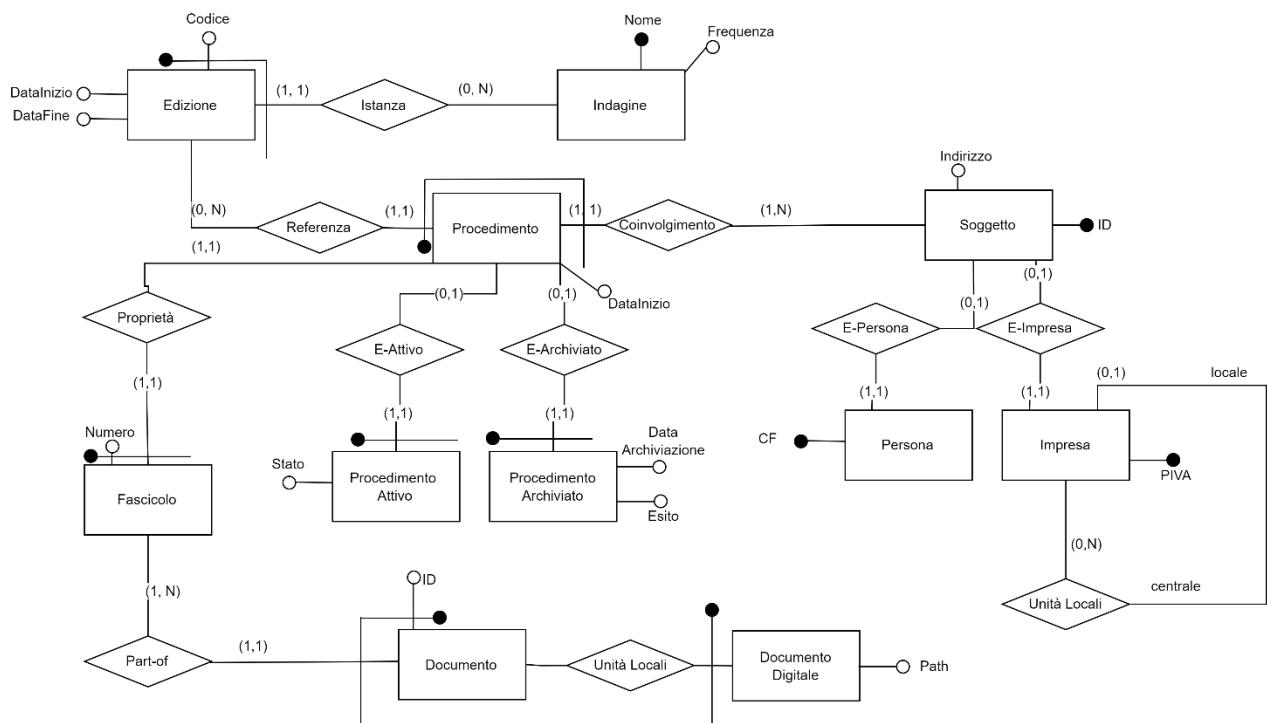


Esercizio 2: Progettazione Logica (7 punti)

A partire dal Diagramma ER dell'Esercizio 1, produrre uno schema relazionale del database nel riquadro sottostante minimizzando i valori nulli. Indicare i vincoli di chiave e gli attributi che ammettono valori nulli. Illustrare come ristrutturare l'ER per essere traducibile in uno schema relazionale.¹

Soluzione Esercizio 2

Il seguente è il diagramma ER ristrutturato per la successiva conversione in uno schema relazionale, inteso come insieme di tabelle e loro vincoli. Nel passaggio, vengono risolte tutte le generalizzazioni ed aggiunti identificatori a tutte le entità. In particolare, viene aggiunto un identificatore a SOGGETTO, PROCEDIMENTO ed INDAGINE. Si assume che per una certa edizione di indagine, un soggetto sia coinvolto in un unico procedimento; inoltre, ogni indagine ha un nome differente. Si noti la soluzione particolare di non aggiungere la relazione come parte dell'identificatore di PERSONA e IMPRESA, giacché le due entità hanno già un proprio identificatore.



Lo schema relazionale è il seguente (non fornito interamente):

INDAGINE(Nome, Frequenza)

EDIZIONE(Codice, Indagine, DataInizio, DataFine)

EDIZIONE.Indagine → INDAGINE.Nome

SOGGETTO(ID, Indirizzo)

PROCEDIMENTO(CodEdizione, Indagine, Soggetto, DataInizio)

PROCEDIMENTO.(CodEdizione, Indagine) → Edizione.(Codice, Indagine)

PERSONA(CF, Soggetto) IMPRESA(PIVA, Soggetto, UnitàCentrale*)

PERSONA.Soggetto → Soggetto.ID IMPRESA.Soggetto → Soggetto.ID

IMPRESA.UnitàCentrale → IMPRESA.PIVA

PROCEDIMENTO_ATTIVO(CodEdizione, Indagine, Soggetto, Stato)

PROCEDIMENTO_ATTIVO.(CodEdizione, Indagine, Soggetto) → PROCEDIMENTO.(CodEdizione, Indagine, Soggetto)

...

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 rosso!).

Nome: _____ Num. Matric.: _____

Esercizio 3: Algebra Relazionale & SQL (7 punti)

Si consideri la seguente base di dati per la partecipazione di delegati a meeting:

Delegato(IdDelegato, Nome, Cognome, Organizzazione)

Partecipa(IdDelegato, IdMeeting)

Meeting(IdMeeting, Citta, Nazione, Data)

- A. Nel riquadro, scrivere una query in Algebra Relazionale che restituisca Nome e Cognome dei delegati che hanno partecipato solo a meeting non italiani (2.5 punti).²

DELEGATI_DA_TENERE = $\pi_{IdDelegato}(\text{DELEGATO}) \setminus \pi_{IdDelegato}(\sigma_{\text{Nazione}='\text{Italia}'}(\text{PARTECIPA} \bowtie \text{MEETING}))$

$\pi_{\text{Nome, Cognome}}(\text{DELEGATI_DA_TENERE} \bowtie \text{DELEGATO})$

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

B. Nel riquadro, scrivere una query in Standard SQL che restituisce, per ogni organizzazione O, il numero di delegati che hanno partecipato a tutti i meeting. (2.5 punti).³

```
CREATE DELEGATI_TUTTI_MEETING (IdDelegato) AS
SELECT IdDelegato
FROM PARTECIPA
GROUP BY IdDelegato
HAVING COUNT(*)=(SELECT COUNT(*) FROM MEETING)

SELECT Organizzazione, COUNT(*)
FROM DELEGATI_TUTTI_MEETING DT, DELEGATO D
WHERE DT.IdDelegato=D.IdDelegato
GROUP BY Organizzazione
```

C. Nel riquadro, scrivere una query in Standard SQL che restituisce, per ogni delegato, la data del primo meeting a cui ha partecipato (2 punti).

```
SELECT P.IdDelegato, MIN(Data)
FROM PARTECIPA P, MEETING M
WHERE P.IdDelegato=M.IdDelegato
GROUP BY M.IdDelegato
```

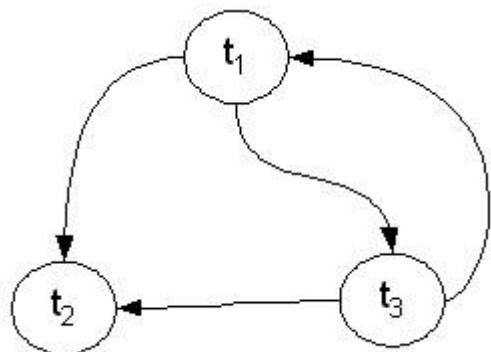
³ Vengono restituite coppie del tipo (O,N) dove O è una organizzazione e N è il numero di delegati di O che hanno partecipato a tutti i meeting

Nome: _____ Num. Matric.: _____

Esercizio 4: Normalizzazione (5 punti)

Dato lo schedule $S = r1(x) \ r1(t) \ r2(z) \ w3(x) \ w1(x) \ r1(y) \ w3(t) \ w2(x) \ w1(y)$, indicare se S è view e/o conflict-serializzabile, motivando la risposta. Se S è view e/o conflict-serializzabile, indicare uno schedule seriale che è view e/o conflict-equivalente, motivando la risposta.

Il grafo dei conflitti è come segue:



Siccome il grafo contiene un ciclo, lo schedule non è conflict-serializzabile.

Tuttavia è view-serializzabile, giacché view-equivalente al seguente schedule seriale:

$$S2 = r1(x) \ r1(t) \ w1(x) \ r1(y) \ w1(y) \ w3(x) \ w3(t) \ r2(z) \ w2(x)$$

Infatti, hanno le stesse relazioni leggi-da e le stesse scritture finali. In particolare, la relazione leggi-data è vuota, mentre le scritture finali sono le seguenti sia per S che per $S2$: $w2(x)$, $w1(y)$ e $w3(t)$

Esercizio 5: Quiz (5 punti)

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

Domanda 1 (1.5 Punti)

Sia data la seguente relazione $R(A, B, C, D)$ con dipendenze funzionali $A \rightarrow D$, $B \rightarrow A, B \rightarrow C$. Indicare quale delle seguenti affermazioni è vera:

1. B è chiave ma **non** è superchiave
2. B è sia chiave che superchiave
3. B **non** è chiave ma è superchiave
4. B **non** è né chiave né superchiave

Domanda 2 (1.5 Punti)

Data la relazione $R(A, B, C, D)$ e $S(W, X, Y, Z)$, indicato con $|R|$ e $|S|$ il numero di tuple di R e S. Sapendo che non ci sono chiavi esterne, quale affermazione è vera per l'operazione $S \bowtie_{A=X} R$ in Algebra Relazionale?

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

Domanda 3 (2 Punti)

Si consideri le relazioni $R(A, B, C, D)$ e la seguente query

SELECT MIN(A) FROM R WHERE B=7

Quale dei seguenti indici garantisce l'efficienza massima?

1. Indice Hash sulla coppia (B,A)
2. Indice Hash sulla coppia (A,B)
3. Indice B-TREE sulla coppia (A,B)
4. Indice B-TREE sulla coppia (B,A)

Nome: _____ Num. Matric.: _____



Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 26 Agosto 2024

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 ripetere il progetto
 - verranno segnalati agli uffici competenti dell'Università

Esercizio 1: Diagramma ER (7 punti)

*Si richiede di progettare lo schema ER concettuale di un'applicazione relativa alle edicole per la vendita di giornali. **Disegnare il diagramma ER nel riquadro della pagina che segue.***

Di ogni edicola interessa il comune in cui essa è registrata, il codice, che è unico nell'ambito del comune in cui l'edicola stessa è registrata, la categoria, l'anno di inizio attività (non sempre disponibile), e i contratti che l'edicola ha con i distributori per l'approvvigionamento dei quotidiani.

Ogni contratto riguarda un'edicola, un quotidiano ed un distributore, ed è caratterizzato dal costo mensile a carico dell'edicola. Infine, per ogni edicola, interessa conoscere le varie persone che sono state proprietarie dell'edicola nei diversi anni, tenendo conto del fatto che in ogni anno un'edicola ha al massimo un proprietario.

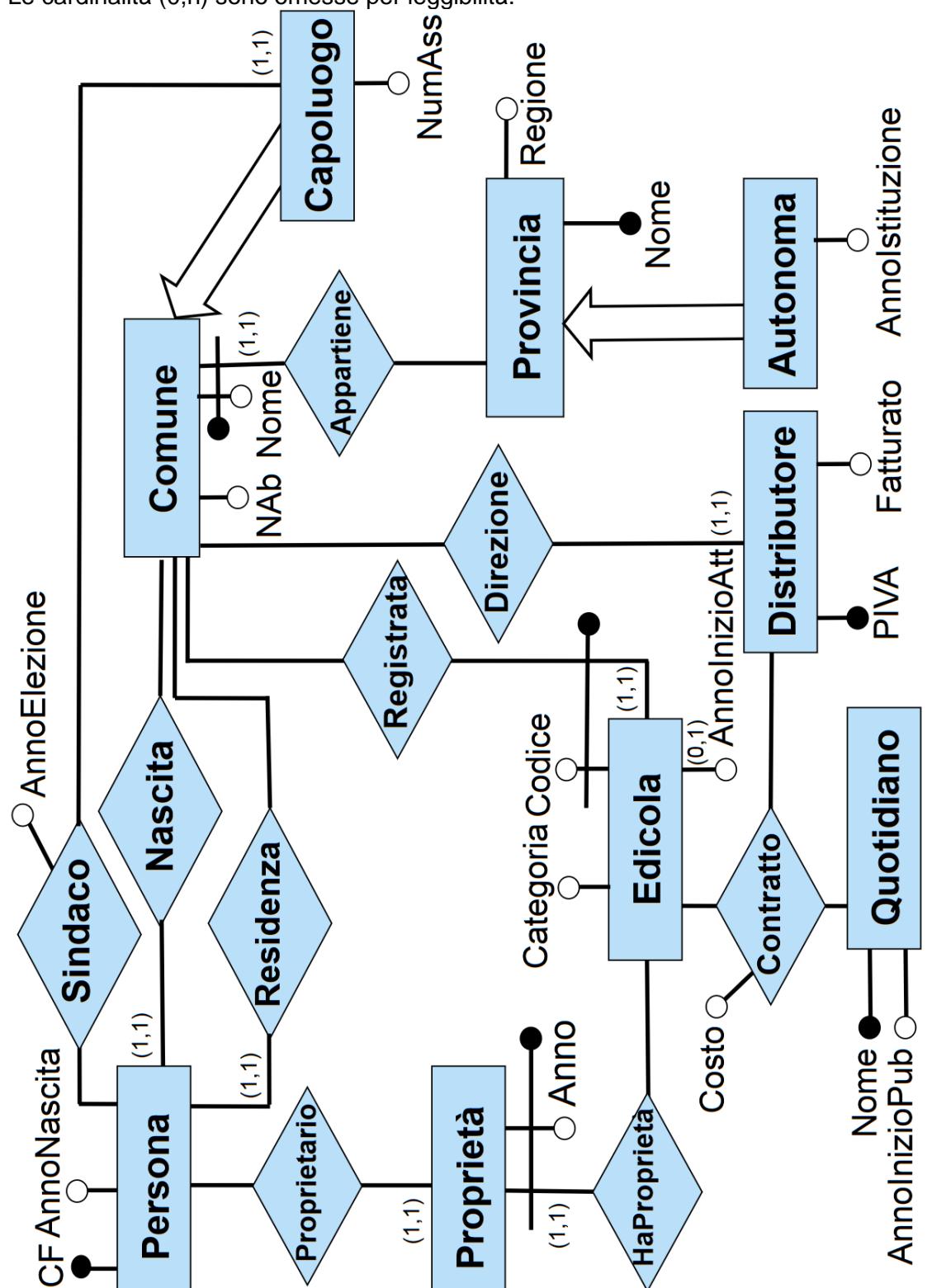
Di ogni persona interessa il codice fiscale (id), l'anno di nascita, il comune di nascita, ed il comune di residenza. Di ogni distributore di quotidiani interessa la partita IVA (id), il fatturato ed il comune in cui è situata la direzione. Di ogni quotidiano interessa il nome (identificativo), e l'anno di inizio pubblicazione.

Di ogni comune interessa la provincia di appartenenza, il nome (unico nella provincia), ed il numero di abitanti. Dei comuni che sono capoluogo di provincia interessa l'attuale sindaco (con l'anno di elezione), ed il numero di assessori comunali. Di ogni provincia interessa il nome (identificativo) e la regione di appartenenza. Alcune province sono “autonome”, e di esse interessa anche l'anno di istituzione.

Nome: _____ Num. Matric.: _____

Soluzione Esercizio 1

Le cardinalità (0,n) sono omesse per leggibilità.



Esercizio 2: Progettazione Logica (7 punti)

A partire dal Diagramma ER dell'Esercizio 1, produrre uno schema relazionale del database nel riquadro sottostante minimizzando il numero di tabelle. Indicare i vincoli di chiave e gli attributi che ammettono valori nulli. Illustrare come ristrutturare l'ER per essere traducibile in uno schema relazionale.¹

Soluzione Esercizio 2

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!)

Nome: _____ Num. Matric.: _____

Esercizio 3: Algebra Relazionale & SQL (7 punti)

Si consideri la seguente base di dati per la gestione di stabilimenti balneari italiani di una catena:

Stabilimento(IdStabilimento, Indirizzo, Città)

Prenotazione(IdStabilimento, CF-Cliente, Data, Costo)

- A. Nel riquadro, scrivere una query in Algebra Relazionale che restituisca i codici fiscali dei clienti che nel 2023 hanno prenotato in esattamente un unico stabilimento della catena (2 punti).²

$$P1 = \sigma_{Data >= 1/1/2023 \text{ AND } Data <= 31/12/2023} (\text{PRENOTAZIONE})$$

$$P2 = P1$$

$$P3CF = \Pi_{P1.CF-Cliente} ($$

$$P1 \bowtie P1.CF-Cliente = P2.CF-Cliente \text{ AND } P1.IdStabilimento <> P2.IdStabilimento \quad P2)$$

$$\Pi_{CF-Cliente} (P1) - P3CF$$

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.

B. Nel riquadro, scrivere una query in Standard SQL che restituisce l'identificativo dello stabilimento che, in media, giornalmente ha più clienti. (2.5 punti).³

```
CREATE CLIENTIXSTABILMENTOXDATA AS
SELECT IdStabilimento, Data, COUNT(*) as Numero
FROM PRENOTAZIONE
GROUP BY IdStabilimento, Data

CREATE NUM_AVG_CLIENTIXSTABILMENTO AS
SELECT IdStabilimento, AVG(Numero) as MediaClienti
FROM CLIENTIXSTABILMENTOXDATA
GROUP BY IdStabilimento

SELECT IdStabilimento
FROM NUM_AVG_CLIENTIXSTABILMENTO
WHERE MediaClienti=
    (SELECT MAX(MediaClienti) FROM
        NUM_AVG_CLIENTIXSTABILMENTO)
```

C. Nel riquadro, scrivere una query in Standard SQL che restituisce, per ogni stabilimento con identificativo S a Rimini, una coppia (S,N) dove N è il numero di clienti diversi che hanno prenotato nello stabilimento con id S (2.5 punti).

```
SELECT P.IdStabilimento, COUNT(DISTINCT CF-Cliente)
FROM STABILIMENTO S, PRENOTAZIONE P
WHERE S.IdStabilimento=P.IdStabilimento AND Citta='Rimini'
GROUP BY P.IdStabilimento
```

³ Qualora ci fossero più stabilimenti che soddisfano i requisiti, devono essere restituiti tutti gli identificativi

Nome: _____ Num. Matric.: _____

Esercizio 4: Ripristino del sistema dopo un guasto (6 punti)

Si supponga che il sistema di gestione di Basi di Dati sia sottoposto ad un guasto di dispositivo (“hard”), in cui si perde la memoria secondaria (per esempio il disco rigido). Sia data la seguente porzione di log fino al gusto:

DUMP, B(T5), B(T6), CK(T5,T6), B(T7), U(T7,O6,B6,A6), U(T6, O3, B7, A7), B(T8), I(T8,O5,A5), C(T8), A(T5).

Indicare il tipo di ripresa da attuare e le operazioni da effettuare per il ripristino del sistema (incluse tutte le letture e scritture).

Occorre effettuare una ripresa a freddo:

1. Il ripristino del sistema alla situazione dell'ultimo DUMP.
2. Occorre ripetere tutte le operazioni dopo il DUMP: B(T5), B(T6), B(T7), U(T7,O6,B6,A6), U(T6, O3, B7, A7), B(T8), I(T8,O5,A5), C(T8), A(T5)
3. Occorre effettuare una ripresa a caldo

Per la ripresa a caldo del punto 3, occorre trovare le transazioni di cui fare REDO e UNDO:

- REDO={T8}
- UNDO={T5, T6, T7}

Vengono a questo punto effettuate le operazioni di UNDO:

1. O3=B7
2. O6=B6

Vengo quindi effettuate tutte le operazioni di REDO:

1. O5=A5

Esercizio 5: Quiz (5 punti)

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

Domanda 1 (1.5 Punti)

Sia data la seguente relazione $R(A, B, C, D)$ con dipendenze funzionali $A \rightarrow BD$, $B \rightarrow A$, $B \rightarrow C$. Indicare quale delle seguenti affermazioni è vera:

1. R è in 3NF ma **non è** in BCNF
2. R è sia in 3NF che in BCNF
3. R non è in 3NF ma è in BCNF
4. R non è né in 3NF né in BCNF

Domanda 2 (1.5 Punti)

Data la relazione $R(A, B, C, D)$, indicato con $|R|$ il numero di tuple di R.

Quante tuple sono presenti nel risultato della seguente operazione in Algebra Relazionale $\sigma_{A='val1' \text{ AND } B='val2'}(R)$?

1. Il numero di tuple è sempre minore di 2;
2. Il numero può essere sia minore di $|R|$, che uguale a $|R|$;
3. Il numero è sempre uguale a $|R|$
4. Il numero è sempre uguale a 0

Domanda 3 (2 Punti)

Si consideri le relazioni $R(A, B, C, D)$ e la seguente query

SELECT AVG(A) FROM R WHERE B=7

Quale dei seguenti indici garantisce l'efficienza massima?

1. Indice Hash sulla coppia (A,B)
2. Indice Hash su B
3. Indice Hash sulla coppia (B,A)
4. Indice B-TREE sulla coppia (B,A)
5. Indice B-Tree su B

Nome: _____ Num. Matric.: _____



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Corso di Laurea in Informatica

Esame di Basi di Dati

Esame del 12 Settembre 2024

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 ripetere il progetto
 - verranno segnalati agli uffici competenti dell'Università

Esercizio 1: Diagramma ER (7 punti)

Si richiede di progettare lo schema concettuale di un'applicazione relativa ai matrimoni civili. Disegnare il diagramma ER nel riquadro della pagina che segue.

I matrimoni sono contratti da due persone. Di ogni persona interessa il codice fiscale (identificativo), l'età, il sesso ed il comune di nascita. Alcune di queste persone officiano i matrimoni (vedasi sotto).

Di ogni matrimonio interessa:

1. il comune e la data in cui è stato celebrato,
2. il codice (unico nell'ambito del comune in cui è stato celebrato),
3. il costo e la sala in cui il matrimonio è ufficiato,
4. il funzionario (di cui interessa l'anzianità) che lo ha officiato,
5. le persone che lo hanno contratto, e
6. gli eventuali figli nati nell'ambito del matrimonio stesso.

Dei comuni interessa il nome (identificativo) ed il numero di abitanti. Delle sale di matrimonio, interessa la dimensione in mq, l'orario di disponibilità, l'indirizzo, il comune in cui è ubicato e la persona responsabile per la sala. Ogni persona può essere responsabile di più sale, ed ogni sala può avere più responsabili. Si vuole anche sapere la data di inizio di responsabilità per una sala, e l'eventuale data di fine (se non è più responsabile per la sala).

Esistono esattamente due tipi di matrimoni: matrimoni validi e matrimoni terminati. Dei matrimoni validi, quelli i cui coniugi sono ancora sposati, interessa il regime finanziario (separazione dei beni, comunione dei beni, misto). Ovviamente, ogni persona è coinvolta al più in un matrimonio valido. Dei matrimoni terminati, quelli finiti per una qualunque causa, interessa la durata espressa in mesi.

Nome: _____

Num. Matric.: _____

Soluzione Esercizio 1

Esercizio 2: Progettazione Logica (7 punti)

A partire dal Diagramma ER dell'Esercizio 1, produrre uno schema relazionale del database nel riquadro sottostante minimizzando il numero di tabelle. Indicare i vincoli di chiave e gli attributi che ammettono valori nulli. Illustrare come ristrutturare l'ER per essere traducibile in uno schema relazionale.¹

Soluzione Esercizio 2

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!)

Nome: _____ Num. Matric.: _____

Esercizio 3: Algebra Relazionale & SQL (7 punti)

Si consideri la seguente base di dati per la gestione dei campionati di calcio di Serie A, dove una stagione è rappresentata dalla stringa Anno-Inizio-AnnoFine (per es. "2023-2024"):

- **Partita(SquadraCasa, SquadraTrasferta, Stagione, GoalCasa, GoalTrasferta)**
- **Luogo(Squadra, Citta)**

A. Nel riquadro, scrivere una query in Algebra Relazionale che restituisca la squadra che ha fatto più goal alla Roma, quando giocava in casa, nella stagione 2023-2024 (2.5 punti).²

P1=

$\sigma_{\text{Stagione}=\text{"2023-2024"} \text{ AND } \text{SquadraTrasferta}=\text{'Roma'}}$ (PARTITA)

P2=P1

P3= $\pi_{\text{SquadraCasa}} (\text{P1} \bowtie \text{P1.GoalCasa} < \text{P2.GoalCasa} \text{ P2})$

$\pi_{\text{SquadraCasa}} (\text{P1}) - \text{P2}$

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.

B. Scrivere una query in SQL Standard che restituisce, per ogni squadra **s** della stagione 2023-2024, la coppia **(s,g)** per indicare che la squadra **s** ha realizzato **g** goal in casa nella stagione. Ordinare poi le coppie per valori di **g** decrescente (2 punti).

```
SELECT SQUADRACASA, SUM(GOALCASA) AS SOMMA_GOAL  
FROM PARTITA  
WHERE STAGIONE="2023-2024"  
GROUP BY SQUADRACASA  
ORDER BY SOMMA_GOAL
```

C. Scrivere una query in SQL Standard che restituisce, per ogni squadra **s** che ha giocato almeno 10 derby come squadra di casa (cioè partite tra squadre della stessa città), il numero medio di goal segnati da **s** nei derby in cui **s** era la squadra di casa (2.5 punti).

```
SELECT L1.SQUADRA, AVG(GOALCASA)  
FROM PARTITA, LUOGO L1, LUOGO L2  
WHERE L1.SQUADRA=SQUADRACASA AND  
L2.SQUADRA=SQUADRATRASFERTA AND  
L1.CITTA=L2.CITTA  
GROUP BY L1.SQUADRA  
HAVING COUNT(L1.SQUADRA)>9
```

Nome: _____ Num. Matric.: _____

Esercizio 4: Normalizzazione (6 punti)

Sia data la relazione R(A,B,C,D,E,F,G,H,I) con copertura ridotta
 $G=\{AB \rightarrow C, AD \rightarrow GH, BD \rightarrow EF, A \rightarrow I\}$

- Trovare la/e chiave/i candidata/e di R, motivando la risposta.
- Quali dipendenze violano la 3NF? Motivare la risposta.
- Effettuare una decomposizione in 3NF.

Punto a

Si considerano le seguenti chiusure

$$AB^+ = \{A, B, C, I\}$$

$$AD^+ = \{A, D, G, H\}$$

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

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

Nessuna delle seguenti chiusure contiene tutti gli attributi. Tuttavia, osservando che l'unione di AB^+ , AD^+ e BD^+ contiene tutti gli attributi, una chiave candidata può essere A,B,D.

Punto b

Tutte le dipendenze violano la 3NF

Punto c

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

$$\{AB \rightarrow C\}, \{AD \rightarrow GH\}, \{BD \rightarrow EF\}, \{A \rightarrow I\}$$

2. Viene costruita una relazione per ogni sottoinsieme:

R1 (A,B,C), R2(A,D,G,H), R3(B,D,E,F), R4(A,I).

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) :

Occorre aggiungere la relazione R5(A, B, D)

5. Indicare le chiavi delle relazioni ottenute dalla normalizzazione

Da completare.

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 A=4 AND B=7` sulla relazione $S(A, B, C)$. Quale dei seguenti indici in genere assicura le migliori performance in termini di velocità dell'esecuzione della query?

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

Domanda 2 (1.5 Punti)

Data la relazione $R(A, B, C, D)$ con dipendenze funzionali $A \rightarrow CD$ e $B \rightarrow A$.

Una delle seguenti affermazioni è vera.

1. AB è chiave e superchiave
2. AB è chiave ma non è superchiave
3. AB non è chiave e ma è superchiave
4. AB non è chiave, né superchiave.

Domanda 3 (2 Punti)

Si consideri lo schedule $S = \{r1(a) w2(a) r3(a) w1(b) w3(c) r2(d) r3(d) w3(b) w3(e)\}$

Dopo aver calcolato il grafo dei conflitti, indicare se S è conflict-serializzabile (CSR) e/o view-serializzabile (VSR)?

1. R è in CSR **ma non** in VSR
2. R **non** è in CSR **ma è** in VSR
3. R è in CSR ed è in VSR
4. R **non** è in CSR **e non** è VSR