

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 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 scaricare il documento stesso.

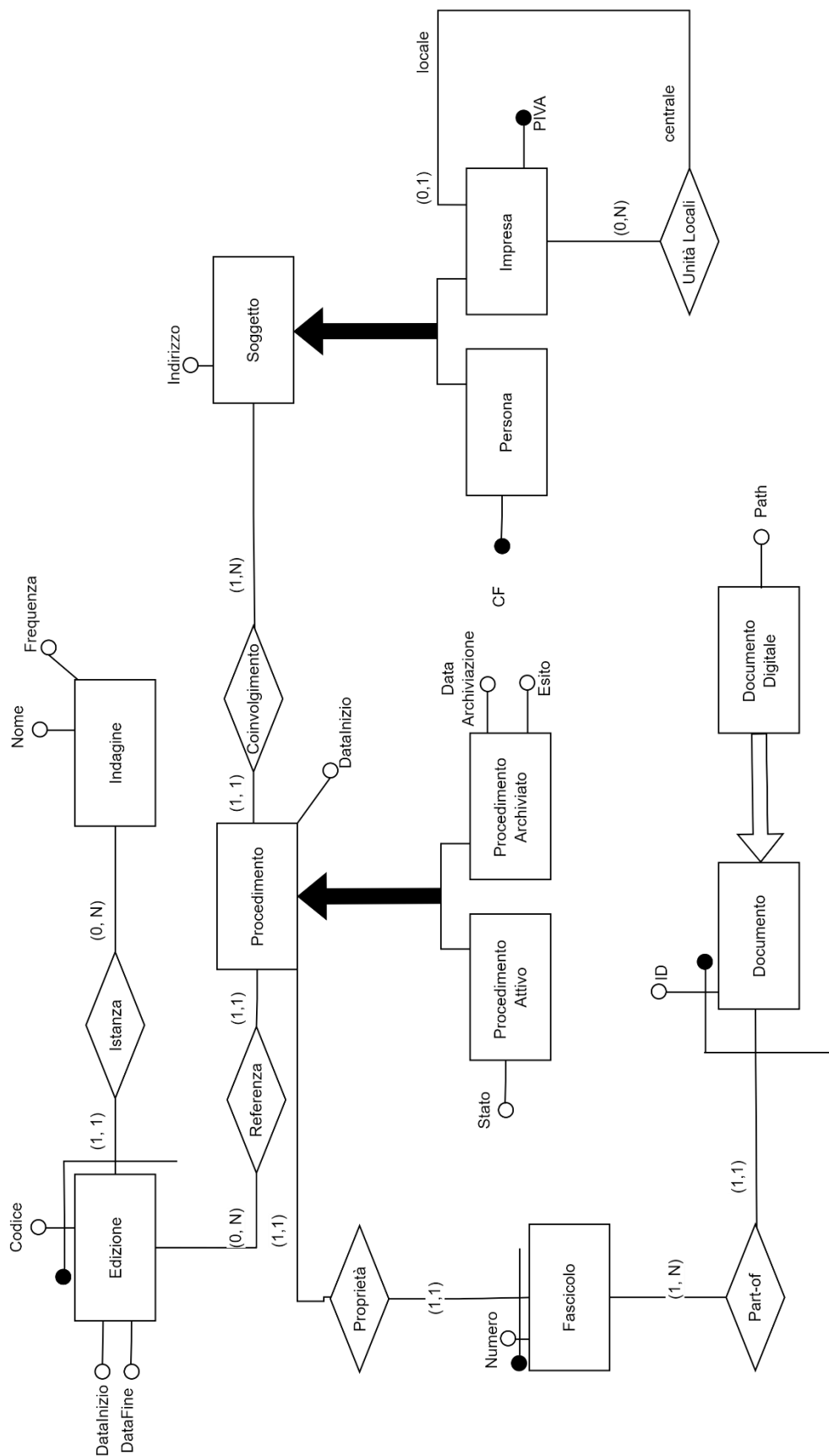
I soggetti possono essere persone fisiche o imprese. Delle persone fisiche interessa memorizzare il codice fiscale, delle imprese la partita iva. Inoltre, per entrambi, è di interesse l'indirizzo cui il soggetto è contattabile.

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 altre 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.

Soluzione Esercizio 1

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

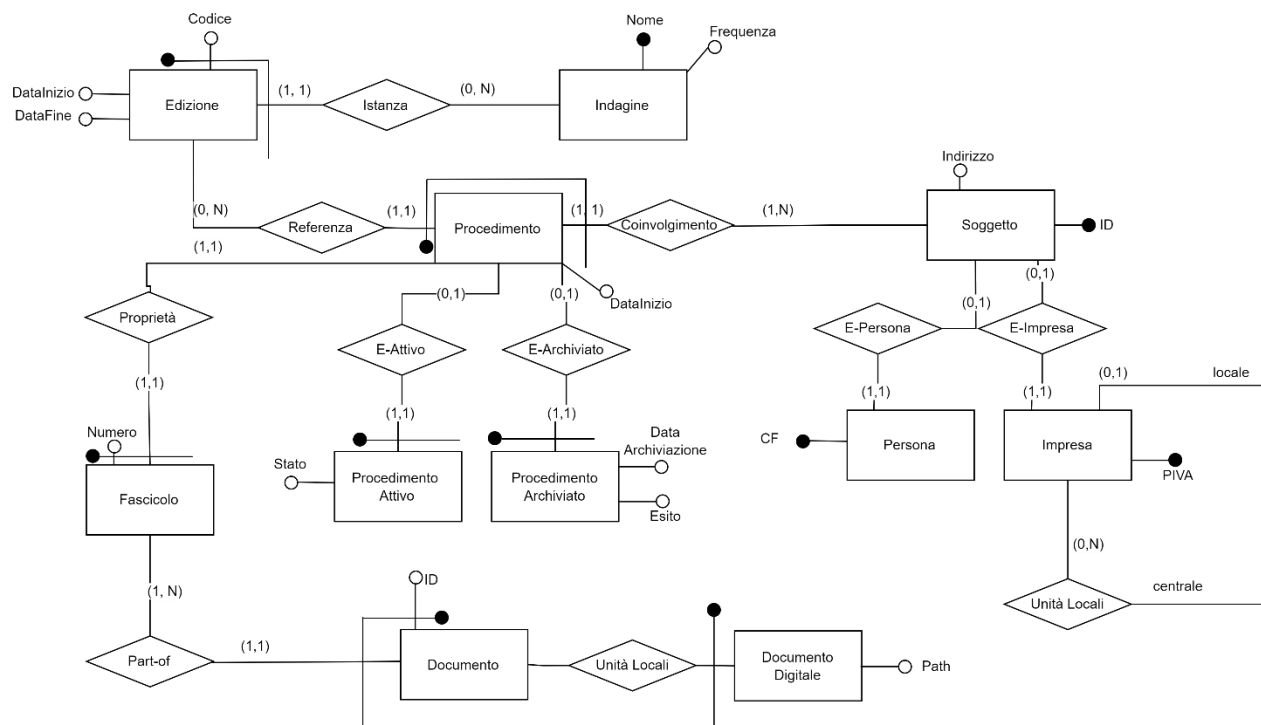


Esercizio 2: Progettazione Logica (7 punti)

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 rossa!)

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

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

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, 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
```

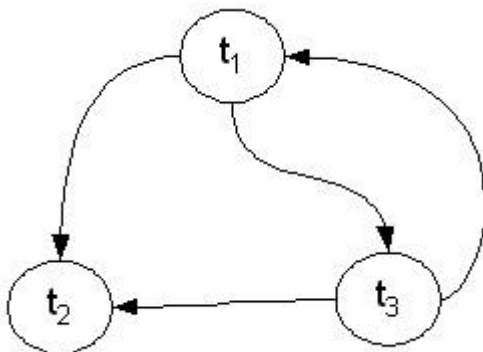
³ Vengono restituite coppie del tipo (O,N) dove O è una organizzazione e N è il numero di delegati di O che hanno partecipato a tutti i meeting

Nome: _____ Num. Matric.: _____

Esercizio 4: Normalizzazione (5 punti)

Dato lo schedule $S = r_1(x) r_1(t) r_2(z) w_3(x) w_1(x) r_1(y) w_3(t) w_2(x) w_1(y)$, 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:

$S_2 = r_1(x) r_1(t) w_1(x) r_1(y) w_1(y) w_3(x) w_3(t) r_2(z) w_2(x)$

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 S_2 : $w_2(x)$, $w_1(y)$ e $w_3(t)$

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 $R(A, \underline{B}, C, D)$ e $S(\underline{W}, X, Y, Z)$, 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)