

# Esercitazione

Hai un database XML che rappresenta una libreria di libri. Ogni libro contiene informazioni come titolo, autore, genere, prezzo e data di pubblicazione. Devi scrivere query XQuery per estrarre informazioni specifiche.

```
<library>
  <book>
    <title>Programming in Python</title>
    <author>John Doe</author>
    <genre>Programming</genre>
    <price>39.99</price>
    <publish_date>2022-05-01</publish_date>
  </book>
  <book>
    <title>XML for Beginners</title>
    <author>Jane Smith</author>
    <genre>Technology</genre>
    <price>29.99</price>
    <publish_date>2021-09-15</publish_date>
  </book>
  <book>
    <title>Advanced XQuery</title>
    <author>James Brown</author>
    <genre>Programming</genre>
    <price>49.99</price>
    <publish_date>2020-11-20</publish_date>
  </book>
  <book>
    <title>Cooking 101</title>
    <author>Mary Johnson</author>
    <genre>Cooking</genre>
    <price>19.99</price>
    <publish_date>2019-07-08</publish_date>
  </book>
</library>
```

- Trova tutti i titoli dei libri di genere "Programming".
- Estrai il nome dell'autore e il titolo di tutti i libri pubblicati dopo il 1° gennaio 2021.
- Calcola il prezzo totale dei libri nella libreria.
- Trova il libro più costoso e visualizza titolo, autore e prezzo.
- Genera una lista di tutti i libri ordinati per data di pubblicazione in ordine decrescente.

# Trova tutti i titoli dei libri di genere "Programming"

```
for $book in doc("library.xml")/library/book
where $book/genre = "Programming"
return $book/title
```

# Estrai il nome dell'autore e il titolo di tutti i libri pubblicati dopo il 1° gennaio 2021

```
for $book in doc("library.xml")/library/book
where xs:date($book/publish_date) >
xs:date("2021-01-01")
return <result>
    <author>{$book/author/text()}</author>
    <title>{$book/title/text()}</title>
</result>
```

# Calcola il prezzo totale dei libri nella libreria.

```
sum(doc("library.xml")/library/book/price)
```

# Trova il libro più costoso e visualizza titolo, autore e prezzo

```
let $maxPrice := max(doc("library.xml")/library/book/price)
for $book in doc("library.xml")/library/book
where $book/price = $maxPrice
return
```

```
    <result>
```

```
        <title>{$book/title/text()}</title>
```

```
        <author>{$book/author/text()}</author>
```

```
        <price>{$book/price/text()}</price>
```

```
    </result>
```

# Genera una lista di tutti i libri ordinati per data di pubblicazione in ordine decrescente

```
for $book in doc("library.xml")/library/book order by
    ($book/publish_date) descending
return
    <book>
        <title>{$book/title/text()}</title>
        <author>{$book/author/text()}</author>
        <publish_date>{$book/publish_date/text()}</publish_date>
    </book>
```

- Si considerino i seguenti schedule:
  - $r_3(y), w_1(y), r_1(x), w_5(x), r_2(z), w_1(t), r_3(z), w_4(x), r_5(t), w_5(z), r_4(t), w_2(z)$
  - $r_2(y), w_2(y), r_3(z), r_1(x), w_1(x), r_3(y), w_3(y), r_2(z), w_3(z), r_2(x), w_2(x), r_1(y), w_1(y).$

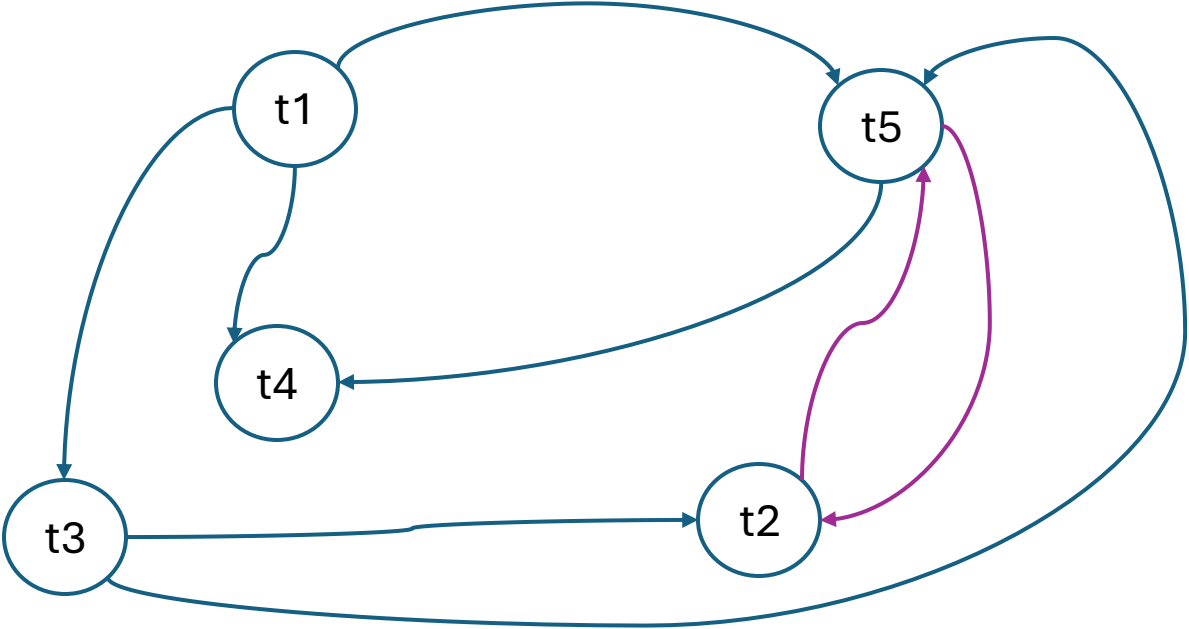
Stabilire se sono VSR o CSR

Se passati ad un 2PL producono deadlock?



$r_3(y), w_1(y), r_1(x), w_5(x), r_2(z), w_1(t), r_3(z), w_4(x), r_5(t), w_5(z), r_4(t), w_2(z)$

x	r1 w5 w4
y	r3 w1
z	r2 r3 w5 w2
t	w1 r5 r4



NO CSR!

$r3(y), w1(y), r1(x), w5(x), r2(z), w1(t), r3(z), w4(x), r5(t), w5(z), r4(t), w2(z)$

$w2(z), w4(x), w1(t), w1(y)$  sono scritture finali

$r5(t)$  legge da  $w1(t)$

$r4(t)$  legge da  $w1(t)$

$t5 < t2$

$t5 < t4$

$t1 < t5 < t4$

$t3 < t1$

$t3 < t2$

$t3 \ t1 \ t5 \ t4 \ t2$

$r3(y), r3(z), \mathbf{w1(y)}, r1(x), \mathbf{w1(t)}, w5(x), r5(t), w5(z), \mathbf{w4(x)}, r4(t), r2(z), \mathbf{w2(z)}$

$r5(t)$  legge da  $w1(t)$

$r4(t)$  legge da  $w1(t)$

$r2(z)$  legge da  $w5(z)$

$t3 \ t1 \ t5 \ t2 \ t4$

$r3(y), r3(z), \mathbf{w1(y)}, r1(x), \mathbf{w1(t)}, w5(x), r5(t), w5(z), r2(z), \mathbf{w2(z)}, \mathbf{w4(x)}, r4(t)$

$r5(t)$  legge da  $w1(t)$

$r4(t)$  legge da  $w1(t)$

$r2(z)$  legge da  $w5(z)$

NON è VSR

r3(y), w1(y), r1(x), w5(x), r2(z), w1(t), r3(z), w4(x), r5(t), w5(z), r4(t), w2(z)

	R_lock	W_lock
x	<del>t1</del>	<del>t5</del> t4
y	<del>t3</del>	<del>t1</del>
z	<del>t2</del> <del>t3</del>	<del>t2</del> <del>t5</del>
t	<del>t5</del> t4	<del>t1</del>

La transazione 1 è in attesa della transazione 3 per un write lock su y

La transazione 3 termina e rilascia tutti i suoi lock e la t1 può essere risvegliata

La transazione 1 riprende e viene bloccata sulla lettura dell'oggetto x perché c'è un w\_lock di t5

La transazione 1 è in attesa della transazione 5 per un write lock su x

La transazione 4 è in attesa della transazione 5 per un write lock su x

La transazione 5 è in attesa della transazione 2 per un write lock su z

La transazione 2 chiede ed ottiene una lock escalation su z e quindi completa la scrittura e rilascia i suoi lock

L'oggetto z non ha più lock e quindi la transazione 5 può essere risvegliata.

La transazione 5 completa e rilascia i suoi lock

Le transazioni 1 e 4 sono risvegliate

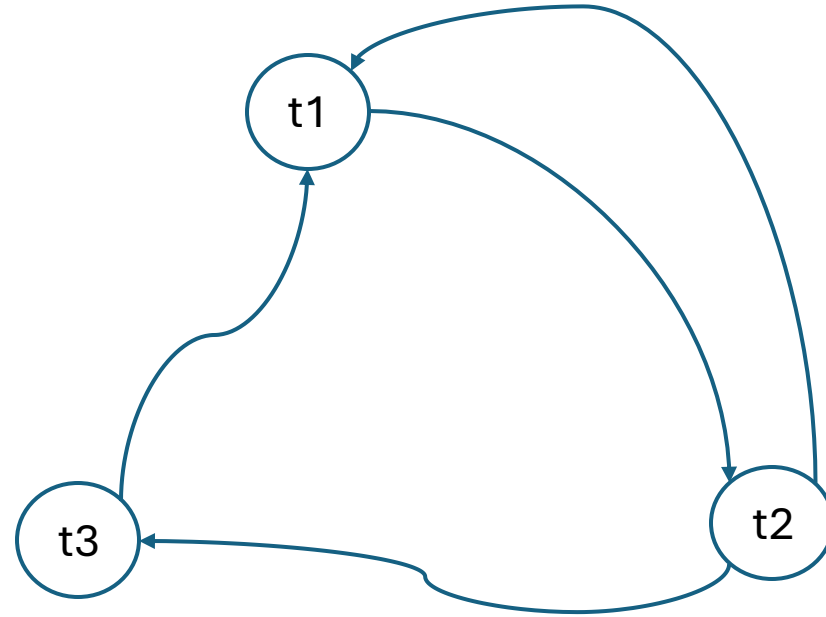
La transazione 1 completa e rilascia i suoi lock

La transazione 4 completa e rilascia i suoi lock

Non si creano deadlock!

$r_2(y), w_2(y), r_3(z), r_1(x), w_1(x), r_3(y), w_3(y), r_2(z), w_3(z), r_2(x), w_2(x), r_1(y), w_1(y)$

x	r1w1r2w2
y	r2w2r3w3r1w1
z	r3r2w3



$r_2(y), w_2(y), r_3(z), r_1(x), w_1(x), r_3(y), w_3(y), r_2(z), \mathbf{w_3(z)}, r_2(x), \mathbf{w_2(x)}, r_1(y), \mathbf{w_1(y)}$

scritture finali:  $w_3(z), w_2(x), w_1(y)$

$r_3(y)$  legge da  $w_2(y)$

$r_2(x)$  legge da  $w_1(x)$

$r_1(x)$  legge da  $w_3(y)$

$t_1: r_1(x), \mathbf{w_1(x)}, r_1(y), \mathbf{w_1(y)}$

$t_2: r_2(y), \mathbf{w_2(y)}, r_2(z), r_2(x), \mathbf{w_2(x)},$

$t_3: r_3(z), r_3(y), \mathbf{w_3(y)}, w_3(z),$

$t_1 < t_2$

$t_3 < t_1$

$t_2 < t_1$

Dato che  $t_2$  deve precedere  $t_1$  affinché  $w_1(y)$  sia scrittura finale, e che  $t_2$  deve anche seguire  $t_1$  affinché  $w_2(x)$  sia scrittura finale, possiamo concludere che lo schedule non può essere VSR in quanto nessuno schedule seriale può soddisfare le condizioni viste prima.

$r_2(y), w_2(y), r_3(z), r_1(x), w_1(x), r_3(y), w_3(y), r_2(z), w_3(z), r_2(x), w_2(x), r_1(y), w_1(y)$

A
A

A
A
A

	R_lock	W_lock
x	<del>t1</del>	t1
y	<del>t2</del>	t2
z	t3 t2	

T2 effettua lock escalation su y

T1 effettua lock escalation su x

T3 vuole leggere y ma non può ottenere r\_lock perché y ha un w\_lock di t2, quindi T3 in attesa di T2

T2 vuole leggere x ma non può ottenere r\_lock perché x ha un w\_lock di t1, quindi T2 in attesa di T1

T1 vuole leggere y ma non può ottenere r\_lock perché y ha un w\_lock di t2, quindi T1 in attesa di T2

Abbiamo un dead lock perché T2 attende una risorsa da T1 e T1 attende una risorsa da T2

Si consideri lo schema relazionale  $R(A,B,C,D,E)$  e il seguente insieme di dipendenze funzionali  $F = \{A \rightarrow B, B \rightarrow C, AC \rightarrow D, D \rightarrow E\}$

- Trovare le chiavi
- Decomporre lo schema in BCNF

- Si Consideri il seguente schema Persona(NOME, CF, NumeroTelefonico, Città) con la seguente dipendenza funzionale  $CF \rightarrow \text{Nome, Città}$ . E' in BCNF (Suggerimento, fare un'istanza d'esempio)?
  - Indicare la chiave primaria e eventualmente decomporlo in BCNF.



- Si consideri lo schema di relazione  $R(A, B, C, D, E, F, G, H, I, J)$  ed il relativo insieme di dipendenze funzionali  $F = \{ABD \rightarrow E, AB \rightarrow G, B \rightarrow F, C \rightarrow J, CJ \rightarrow I, G \rightarrow H\}$ .
  - Stabilire se  $F$  è una copertura minimale. In caso di risposta negativa, determinare una copertura minimale di  $F$ .
  - Determinare l'insieme delle chiavi di  $R$ .

- Siano dati lo schema relazionale  $R(A, B, C, D, E, F)$  e gli insiemi di dipendenze funzionali  $G = \{AB \rightarrow C, B \rightarrow A, AD \rightarrow E, BD \rightarrow F\}$  ed  $H = \{AB \rightarrow C, B \rightarrow A, AD \rightarrow EF\}$ 
  - Determinare una copertura minimale per  $G$  ed una copertura minimale per  $H$ .
  - Stabilire se  $G$  ed  $H$  sono equivalenti.

- Si considerino lo schema di relazione  $R(A, B, C, D, E, F)$  e l'insieme di dipendenze associato:  $F = \{A \rightarrow B, C \rightarrow AD, AF \rightarrow EC\}$ .
  - Si determinino le chiavi candidate di  $R$ .
  - Si stabilisca se  $R$  è in 3NF. Qualora non lo sia, si definisca una decomposizione di  $R$  in 3NF che conservi le dipendenze date.

Si considerino lo schema di relazione  $R(A, B, C, D, E, G)$  e l'insieme di dipendenze associato:  $F = \{E \rightarrow D, C \rightarrow B, CE \rightarrow G, B \rightarrow A\}$ .

- Si stabilisca se  $R$  è in 3NF. Qualora non lo sia, si definisca una decomposizione di  $R$  in 3NF che conservi le dipendenze date.
- Se  $R$  non è in BCNF, determinare una decomposizione di  $R$  in BCNF

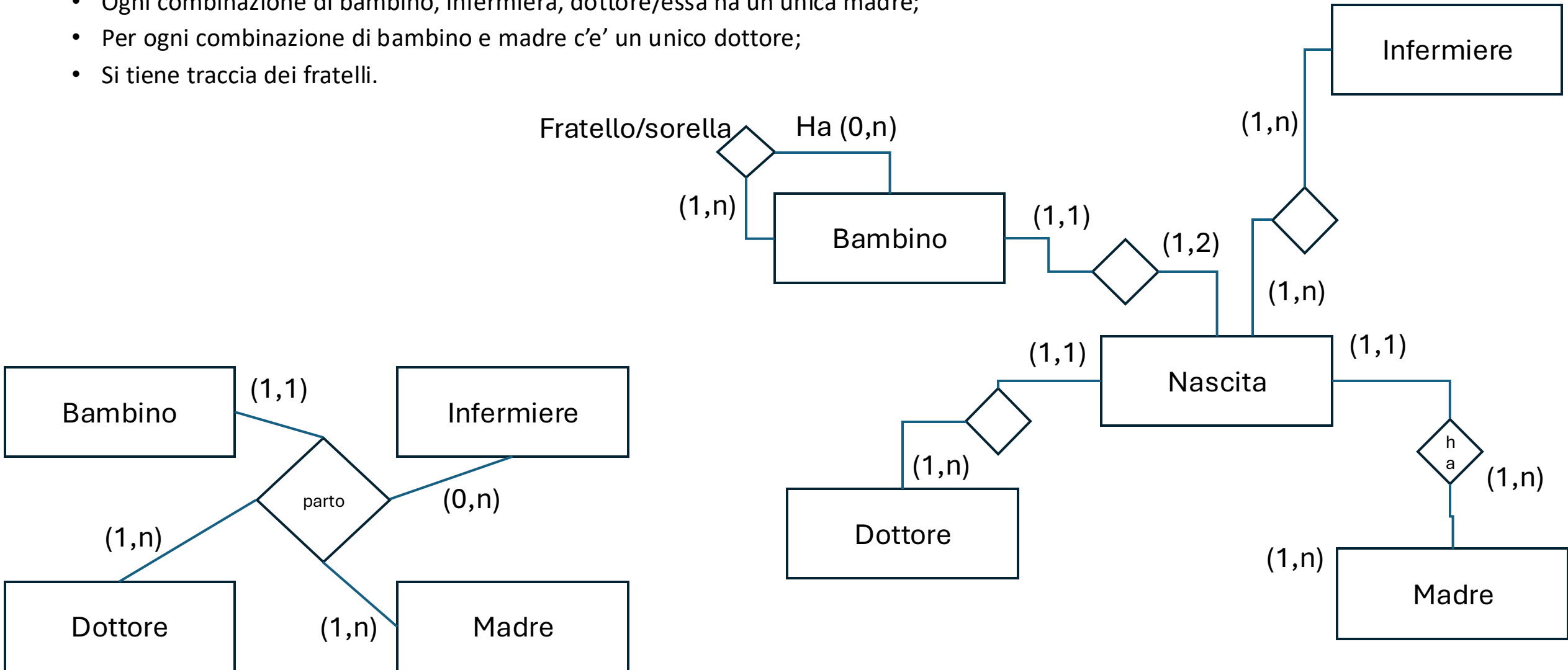
- Si faccia il design di un database che consenta di descrivere le nascite. Alla nascita si ha un bambino (i gemelli sono rappresentati da due nascite), una madre, sono presenti dottori, e infermieri.
1. Rappresentare tutto con 4 entità e una associazione e indicare le opportune cardinalità dell'associazione.
  2. Ridisegnare lo schema supponendo di aggiungere i seguenti vincoli:
    1. Ogni bambino ha un'unica madre;
    2. Ogni combinazione di bambino, infermiera, dottore/essa ha un'unica madre;
    3. Per ogni combinazione di bambino e madre c'è un unico dottore;
    4. Si tiene traccia dei fratelli.

- Si faccia il design di un database che consenta di descrivere le nascite. Alla nascita si ha un bambino (i gemelli sono rappresentati da due nascite), una madre, sono presenti dottori, e infermieri.

1. Rappresentare tutto con 4 entità e una associazione e indicare le opportune cardinalità dell'associazione.

2. Ridisegnare lo schema supponendo di aggiungere i seguenti vincoli:

- Ogni bambino ha un'unica madre;
- Ogni combinazione di bambino, infermiera, dottore/essa ha un'unica madre;
- Per ogni combinazione di bambino e madre c'è un unico dottore;
- Si tiene traccia dei fratelli.



Si vuole progettare una base di dati di supporto alla gestione della rete di stazioni per il rifornimento di carburante presenti sul territorio della regione Sicilia.

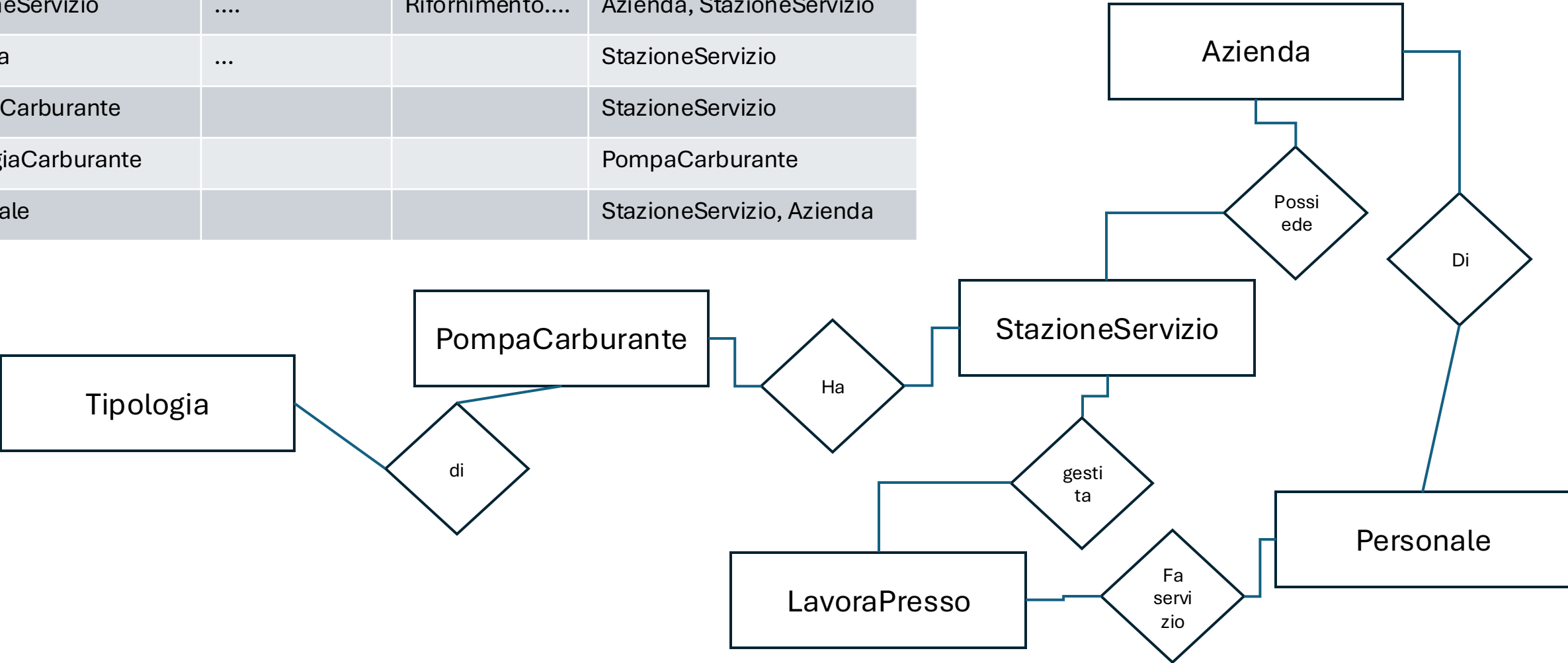
Ogni stazione di rifornimento è identificata univocamente da un codice ed è caratterizzata dall'azienda che la possiede, dalle coordinate geografiche, che identificano la sua posizione, e dal comune di appartenenza. Ogni stazione offre diversi tipi di carburante (benzina, gasolio, gpl, metano,...) e dispone di un certo numero di pompe per l'erogazione del carburante. Assumiamo che non necessariamente ogni stazione offra tutti i tipi possibili di carburante. Si vuole tener traccia delle (poche) stazioni che erogano gas. Per ogni tipo di carburante disponibile presso una data stazione, si memorizzi la capacità massima del relativo serbatoio e la quantità correntemente disponibile. All'interno di una determinata stazione, ogni pompa è caratterizzata da un numero (pompa numero 1, pompa numero 2, ..) ed è caratterizzata dal tipo di carburante erogato.

Termine	Descrizione	Sinonimi	Termini collegati
StazioneServizio	....	Rifornimento....	Azienda StazioneServizio
Azienda	...		StazioneServizio
PompeCarburante			StazioneServizio
TipologiaCarburante			PompaCarburante
Personale			StazioneServizio Azienda

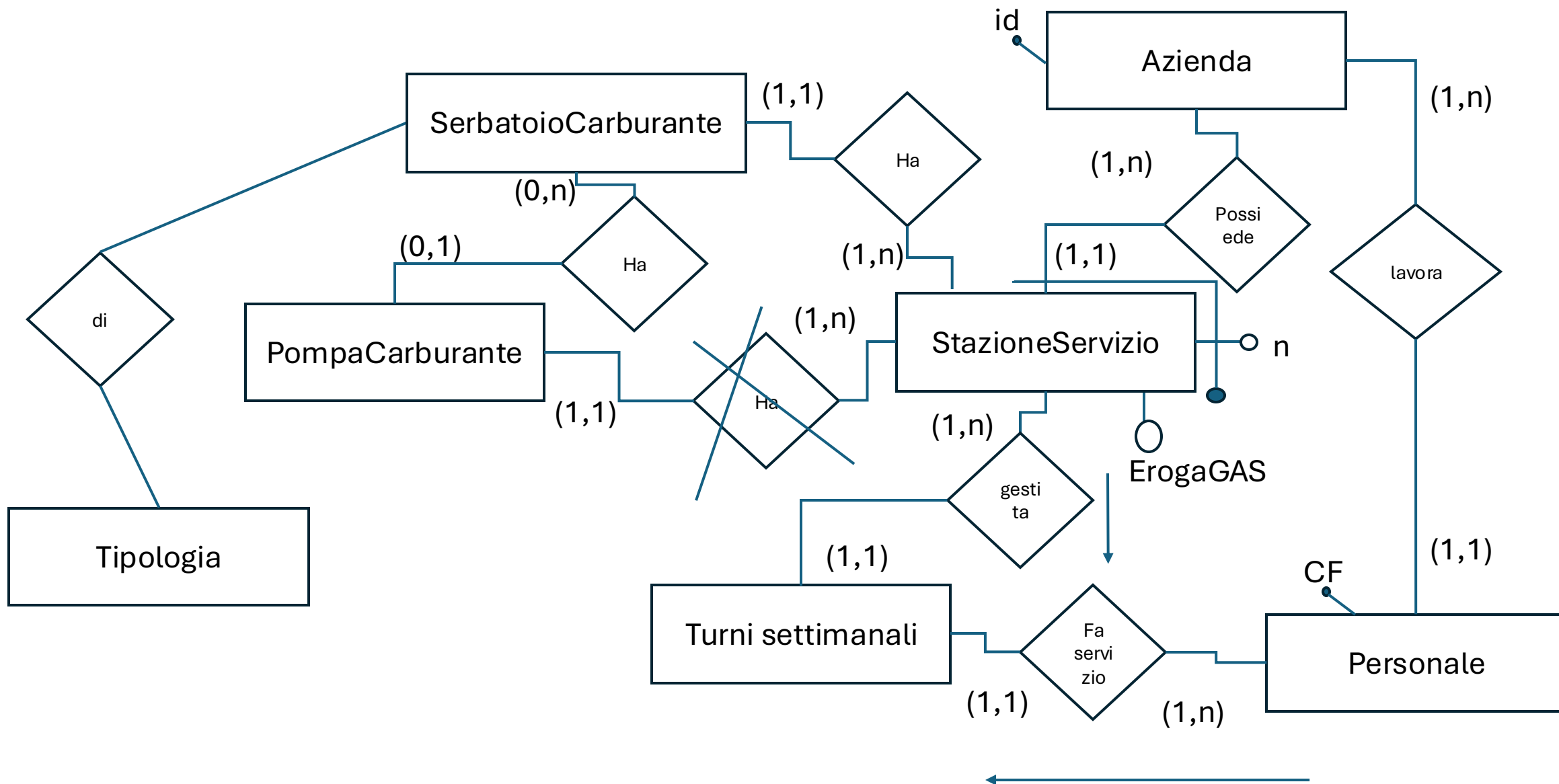
Ogni azienda è identificata univocamente da un codice. Di ogni azienda vogliamo memorizzare il responsabile per la regione Sicilia e il comune dove si trova l'ufficio regionale di riferimento. Si assuma che un'azienda possa possedere più stazioni di rifornimento e che ogni stazione di rifornimento appartenga ad un'unica azienda. Ogni azienda dispone di un certo numero di persone che prestano servizio presso le stazioni di rifornimento di sua proprietà. Ogni dipendente di un'azienda è identificato univocamente dal suo codice fiscale ed è caratterizzato da nome e cognome, residenza e recapito telefonico. Si assuma che un dipendente possa lavorare presso più stazioni di rifornimento dell'azienda. Di ogni dipendente che lavora presso più stazioni, vogliamo memorizzare il piano di lavoro settimanale (si assuma che un dipendente non possa spostarsi da una stazione all'altra in uno stesso giorno della settimana, ma possa lavorare presso stazioni diverse in giorni diversi).

Si definisca uno schema ER che descriva il database, illustrando con chiarezza le eventuali assunzioni fatte.

Termine	Descrizione	Sinonimi	Termini collegati
StazioneServizio	....	Rifornimento....	Azienda, StazioneServizio
Azienda	...		StazioneServizio
PompeCarburante			StazioneServizio
TipologiaCarburante			PompaCarburante
Personale			StazioneServizio, Azienda







- Si vuole progettare una base di dati di supporto alla gestione della rete di stazioni per il rifornimento di carburante presenti sul territorio della regione Sicilia.
- Ogni stazione di rifornimento è identificata univocamente da un codice ed è caratterizzata dall'azienda che la possiede, dalle coordinate geografiche, che identificano la sua posizione, e dal comune di appartenenza. Ogni stazione offre diversi tipi di carburante (benzina, gasolio, gpl, metano,...) e dispone di un certo numero di pompe per l'erogazione del carburante. Assumiamo che non necessariamente ogni stazione offra tutti i tipi possibili di carburante. Si vuole tener traccia delle (poche) stazioni che erogano gas. Per ogni tipo di carburante disponibile presso una data stazione, si memorizzi la capacità massima del relativo serbatoio e la quantità correntemente disponibile. All'interno di una determinata stazione, ogni pompa è caratterizzata da un numero (pompa numero 1, pompa numero 2, ..) e sia caratterizzata dal tipo di carburante erogato.

- Ogni azienda è identificata univocamente da un codice. Di ogni azienda vogliamo memorizzare il responsabile per la regione Sicilia e il comune dove si trova l'ufficio regionale di riferimento. Si assuma che un'azienda possa possedere più stazioni di rifornimento e che ogni stazione di rifornimento appartenga ad un'unica azienda. Ogni azienda dispone di un certo numero di persone che prestano servizio presso le stazioni di rifornimento di sua proprietà. Ogni dipendente di un'azienda è identificato univocamente dal suo codice fiscale ed è caratterizzato da nome e cognome, residenza e recapito telefonico. Si assuma che un dipendente possa lavorare presso più stazioni di rifornimento dell'azienda. Di ogni dipendente che lavora presso più stazioni, vogliamo memorizzare il piano di lavoro settimanale (si assuma che un dipendente non possa spostarsi da una stazione all'altra in uno stesso giorno della settimana, ma possa lavorare presso stazioni diverse in giorni diversi).
- Si definisca uno schema ER che descriva il database, illustrando con chiarezza le eventuali assunzioni fatte.

- Sia dato il seguente insieme di requisiti relativi ad una base di dati per l'archiviazione delle maggiori opere d'arte di artisti italiani presenti nei musei italiani.
  - Ogni museo è identificato univocamente dal nome ed è caratterizzato dal direttore, dalla città dove si trova e dall'indirizzo. Di ogni museo vogliamo, inoltre, memorizzare il numero di opere di artisti italiani presenti.
  - Ogni opera è identificata univocamente, all'interno del museo dove si trova, da un codice numerico (non si escluda la possibilità che uno stesso codice venga associato a due o più opere collocate in musei distinti) e sia caratterizzata dall'artista che l'ha realizzata, dall'anno in cui è stata completata, dal titolo e dal valore commerciale.
  - Ogni artista autore di almeno un'opera presente nella rete dei musei italiani è identificato univocamente dal codice fiscale ed è caratterizzato dal nome, dal cognome, dalla città in cui è nato e dall'anno di nascita. Di ogni artista vogliamo memorizzare il numero di opere presenti nei musei italiani.
  - Si tenga, infine, traccia delle città italiane che sono sede di uno o più musei e/o luoghi di nascita di uno o più artisti.

Si definisca uno schema ER che descriva il contenuto informativo del sistema,

- Immagina di dover progettare una base di dati per gestire il sistema informatico di una scuola. L'obiettivo è sviluppare un sistema che permetta di monitorare e gestire le informazioni relative a studenti, corsi, insegnanti, aule e le iscrizioni ai corsi. Gli studenti, infatti, si iscrivono a uno o più corsi, che sono insegnati da insegnanti specifici e si svolgono in aule con una capacità limitata di posti. L'intero sistema deve essere strutturato per garantire una gestione chiara e ben organizzata dei dati relativi a queste entità.

- Sia dato un file ordinato con  $r = 200000$  record memorizzati su un disco con dimensioni del blocco  $B = 2056$  byte. I record abbiano lunghezza fissa  $R = 400$  byte e siano di tipo unspanned. La dimensione del campo chiave primaria sia  $V1 = 20$  byte, mentre la dimensione del campo chiave candidata (diversa dalla chiave primaria) sia  $V2 = 15$  byte. La dimensione del puntatore a blocco sia  $P = 5$  byte.
  1. Si assuma che il file sia ordinato rispetto al campo chiave primaria. Determinare il numero di accessi a blocco richiesti rispettivamente da una ricerca con indice primario e da una ricerca con indice secondario sulla chiave candidata (diversa dalla chiave primaria).
  2. Si assuma che il file non sia ordinato rispetto ad alcun campo. Determinare il numero di accessi a blocco richiesti rispettivamente da una ricerca con indice secondario sulla chiave primaria e da una ricerca con indice secondario sulla chiave candidata (diversa dalla chiave primaria).

- Sia dato un file ordinato con  $r = 200000$  record memorizzati su un disco con dimensioni del blocco  $B = 2056$  byte. I record abbiano lunghezza fissa  $R = 400$  byte e siano di tipo unspanned. La dimensione del campo chiave primaria sia  $V1 = 20$  byte, mentre la dimensione del campo chiave candidata (diversa dalla chiave primaria) sia  $V2 = 15$  byte. La dimensione del puntatore a blocco sia  $P = 5$  byte.

1. Si assuma che il file sia ordinato rispetto al campo chiave primaria. Determinare il numero di accessi a blocco richiesti rispettivamente da una ricerca con indice primario e da una ricerca con indice secondario sulla chiave candidata (diversa dalla chiave primaria).

- Sia dato un file ordinato contenente  $r = 100000$  record memorizzati su un disco con dimensioni del blocco  $B = 1024$  byte. I record hanno lunghezza fissa  $R = 200$  byte e sono di tipo unspanned. La dimensione del campo chiave sia  $V1 = 10$  byte. Il puntatore ad un blocco sia lungo  $P = 5$  byte.
  1. Si assuma che il file sia ordinato rispetto al campo chiave. Determinare il numero di accessi a blocco richiesti rispettivamente da una ricerca binaria e da una ricerca con indice primario.
  2. Si assuma che il file sia ordinato rispetto ad un campo non chiave di dimensione  $V2 = 20$  byte. Determinare il numero di accessi a blocco richiesti rispettivamente da una ricerca lineare e da una ricerca con indice secondario denso costruito su un campo chiave di dimensione  $V3 = 15$  byte.
  3. Si consideri l'indice secondario denso definito al punto precedente e lo si converta in un indice multilivello. Calcolare il numero di livelli dell'indice, il numero di blocchi per ogni livello ed il numero di accessi necessario per recuperare un blocco del file di dati.