

Basi di dati — 25 gennaio 2018 — Compito A

Tempo a disposizione: un'ora per la prova breve, due ore per la prova lunga

Cognome: _____ Nome: _____ Matricola: _____

Basi di dati I — 25 gennaio 2018 — Compito A

Domanda 1 (50% per le prova breve e 25% per la prova completa)

Considerare la relazione seguente

Cod	Titolo	CC	Città	CN	Nazione	CM	Materia	CA	Area
1	Articolazioni	C1	Milano	N1	Italia	M1	Ortopedia	A1	Medicina
2	Fratture	C2	Roma	N1	Italia	M1	Ortopedia	A1	Medicina
3	Cuore2018	C1	Milano	N1	Italia	M2	Cardiologia	A1	Medicina
4	DB2018	C3	Monaco	N2	Germania	M1	Basi di dati	A2	Informatica

La relazione mostra (in forma non normalizzata) i dati di interesse in un certo anno per un insieme di congressi scientifici, secondo le seguenti specifiche:

- ogni congresso ha un codice e un titolo e si tiene una città
- ogni città ha un nome, si trova in una nazione e ha un codice che la identifica univocamente
- ogni nazione ha un codice e un nome
- ogni congresso è relativo ad una materia specifica (ad esempio l'ortopedia) che a sua volta appartiene ad un'area scientifica (ad esempio la medicina); ogni materia ha un nome e un codice che la identifica nell'ambito dell'area scientifica
- ogni area scientifica ha un codice e un nome

Con riferimento alle specifiche e ai dati nell'esempio

1. mostrare uno schema concettuale per la realtà di interesse (rispettando le specifiche per i vari codici)

2. mostrare una buona decomposizione della relazione originaria che soddisfi la BCNF (mostrare le tabelle, indicando la chiave di ciascuna con la sottolineatura)

Domanda 2 (50% per le prova breve e 25% per la prova completa)

Mostrare uno schema concettuale ancora per una realtà relativa a congressi, ma con le specifiche seguenti, che estendono quelle della domanda precedente:

- interessano congressi svolti anche in anni diversi
- ogni congresso appartiene ad una serie, e per ogni serie (ad esempio, la serie VLDB) esiste al più un congresso all'anno (quindi ad esempio esiste il VLDB 2017); ogni serie ha un codice e un titolo ed è relativa ad una materia (che ha codice, nome e area scientifica con gli stessi dettagli della domanda precedente)
- ogni congresso si svolge in una città (che ha codice, nome e nazione con gli stessi dettagli della domanda precedente), in una certa data (per semplicità un solo giorno)
- ogni congresso è composto di una serie di sessioni, ognuna con codice (univoco nella base di dati), orario e un insieme di partecipanti, ognuno dei quali presenta una relazione con un titolo (semplice stringa di testo) e ha una affiliazione (cioè una istituzione, ad esempio una università o una azienda); si noti che l'affiliazione di una stessa persona può variare da un congresso all'altro
- ogni istituzione ha un codice un nome e una città in cui ha sede

Domanda 3 (30%) Considerare la seguente porzione di base di dati che contiene alcune informazioni su corsi universitari, esami e piani di studio (ovviamente tanto le relazioni quanto gli attributi e anche le ennuple sono solo una parte di quelli di interesse; in particolare, la relazione piani di studio contiene le coppie studente-corso per le quali lo studente ha il corso nel proprio piano di studio)

CORSI		
<u>Codice</u>	Titolo	CFU
101	Fondamenti	12
102	POO	9
103	Basi di dati	6
104	Mobile Comp	6

ESAMI		
<u>Corso</u>	<u>Studente</u>	Voto
101	100001	30
101	100002	28
101	100003	26
102	100001	30
102	100002	27
102	100003	30
103	100001	30
104	100003	28

PIANIDISTUDIO	
<u>Corso</u>	<u>Studente</u>
101	100001
101	100002
101	100003
101	100004
102	100001
102	100002
102	100003
102	100004
103	100001
103	100002
104	100003
104	100004

1. Mostrare il risultato della seguente interrogazione

```
SELECT corso, titolo, CFU, count(studente) AS numeroesamiprevisti
FROM corsi join pianidistudio on codice=corso
GROUP BY corso, titolo, CFU;
```

2. Formulare in SQL l'interrogazione che conta, per ogni corso, quanti studenti lo hanno in piano di studio e non hanno ancora superato l'esame

3. Mostrare il risultato della seguente interrogazione

```
SELECT codice, titolo, CFU,
       count(e.studente) AS numeroesamiprevisti, count(e.studente) AS numeroesamisuperati
FROM corsi join pianidistudio p on codice=p.corso
       join esami e on codice=e.corso
GROUP BY codice, titolo, CFU;
```

4. Formulare in SQL l'interrogazione che restituisce, per ciascun corso, il codice, il titolo, il numero di studenti che lo hanno in piano di studi e il numero di studenti che hanno superato l'esame (cioè ciò che probabilmente si desiderava ottenere con l'interrogazione precedente)

Risultato desiderato:

	codice integer	titolo character (20)	cfu integer	numeroesamiprevisti bigint	numeroesamisuperati bigint
1	101	Fondamenti	12	4	3
2	102	POO	9	4	3
3	103	Basi di dati	6	2	1
4	104	Mobile Comp	6	2	1

Domanda 4 (20%) Lo schema concettuale seguente rappresenta un insieme di corsi e un insieme di iscrizioni a tali corsi. Nello schema l'attributo *Ricavo* è ridondante perché può essere ottenuto moltiplicando il costo del corso per il numero di iscritti (cioè il prodotto del valore dell'attributo *Costo* di ogni occorrenza dell'entità *Corso* per il numero di occorrenze dell'entità *Persona* a cui essa è correlata tramite la relationship *Iscrizione*).



Valutare se convenga o meno mantenere la ridondanza, assumendo che *Corso* abbia $C_C = 1.000$ occorrenze, *Persona* ne abbia $C_P = 3.000$ e *Iscrizione* $C_I = 30.000$ (e quindi che ogni corso abbia mediamente $C_{CI} = 30$ iscritti) e che il carico applicativo includa come operazioni principali le seguenti:

1. calcolo del ricavo di un corso (dato il codice), con frequenza $f_1 = 50$
2. iscrizione di una persona (già nella base di dati e di cui è dato il codice fiscale) ad un corso (di cui è dato il codice), con frequenza $f_2 = 20.000$; in sostanza, questa operazione richiede l'inserimento di una occorrenza di *Iscrizione*

Considerare i costi delle letture e scritture delle occorrenze di entità e relationship e considerare il costo delle scritture pari a quello delle letture.

Basi di dati — 25 gennaio 2018 — Compito B

Tempo a disposizione: un'ora per la prova breve, due ore per la prova lunga

Cognome: _____ Nome: _____ Matricola: _____

Basi di dati I — 25 gennaio 2018 — Compito B

Domanda 1 (50% per le prova breve e 25% per la prova completa)

Considerare la relazione seguente

Cod	Titolo	CC	Città	CN	Nazione	CM	Materia	CA	Area
1	Articolazioni	C1	Milano	N1	Italia	M1	Ortopedia	A1	Medicina
2	Fratture	C2	Roma	N1	Italia	M1	Ortopedia	A1	Medicina
3	Cuore2018	C1	Milano	N1	Italia	M2	Cardiologia	A1	Medicina
4	DB2018	C1	Monaco	N2	Germania	M3	Basi di dati	A2	Informatica

La relazione mostra (in forma non normalizzata) i dati di interesse in un certo anno per un insieme di convegni scientifici, secondo le seguenti specifiche:

- ogni convegno ha un codice e un titolo e si tiene una città
- ogni città ha un nome, si trova in una nazione e ha un codice che la identifica nell'ambito della nazione
- ogni nazione ha un codice e un nome
- ogni convegno è relativo ad una materia specifica (ad esempio l'ortopedia) che a sua volta appartiene ad un'area scientifica (ad esempio la medicina); ogni materia ha un nome e un codice che la identifica univocamente
- ogni area scientifica ha un codice e un nome

Con riferimento alle specifiche e ai dati nell'esempio

1. mostrare uno schema concettuale per la realtà di interesse (rispettando le specifiche per i vari codici)

2. mostrare una buona decomposizione della relazione originaria che soddisfi la BCNF (mostrare le tabelle, indicando la chiave di ciascuna con la sottolineatura)

Domanda 2 (50% per le prova breve e 25% per la prova completa)

Mostrare uno schema concettuale ancora per una realtà relativa a convegni, ma con le specifiche seguenti, che estendono quelle della domanda precedente:

- interessano convegni svolti anche in anni diversi
- ogni convegno appartiene ad una serie, e per ogni serie (ad esempio, la serie VLDB) esiste al più un convegno all'anno (quindi ad esempio esiste il VLDB 2017); ogni serie ha un codice e un titolo ed è relativa ad una materia (che ha codice, nome e area scientifica con gli stessi dettagli della domanda precedente)
- ogni convegno si svolge in una città (che ha codice, nome e nazione con gli stessi dettagli della domanda precedente), in una certa data (per semplicità un solo giorno)
- ogni convegno è composto di una serie di sessioni, ognuna con codice (univoco nella base di dati), orario e un insieme di partecipanti, ognuno dei quali presenta una relazione con un titolo (semplice stringa di testo) e ha una affiliazione (cioè una istituzione, ad esempio una università o una azienda); si noti che l'affiliazione di una stessa persona può variare da un convegno all'altro
- ogni istituzione ha un codice un nome e una città in cui ha sede

Domanda 3 (30%) Considerare la seguente porzione di base di dati che contiene alcune informazioni su corsi universitari, esami e piani di studio (ovviamente tanto le relazioni quanto gli attributi e anche le ennuple sono solo una parte di quelli di interesse; in particolare, la relazione piani di studio contiene le coppie studente-corso per le quali lo studente ha il corso nel proprio piano di studio)

CORSI		
<u>Codice</u>	Titolo	CFU
101	Fondamenti	12
102	POO	9
103	Basi di dati	6
104	Mobile Comp	6

ESAMI		
<u>Corso</u>	<u>Studente</u>	Voto
101	100001	30
101	100002	28
101	100003	26
102	100001	30
102	100002	27
102	100003	30
103	100001	30
104	100003	28

PIANIDISTUDIO	
<u>Corso</u>	<u>Studente</u>
101	100001
101	100002
101	100003
101	100004
102	100001
102	100002
102	100003
102	100004
103	100001
103	100002
104	100003
104	100004

1. Mostrare il risultato della seguente interrogazione

```
SELECT corso, titolo, CFU, count(studente) AS numeroesamiprevisti
FROM corsi join pianidistudio on codice=corso
GROUP BY corso, titolo, CFU;
```

2. Formulare in SQL l'interrogazione che conta, per ogni corso, quanti studenti lo hanno in piano di studio e non hanno ancora superato l'esame

3. Mostrare il risultato della seguente interrogazione

```
SELECT codice, titolo, CFU,
       count(e.studente) AS numeroesamiprevisti, count(e.studente) AS numeroesamisuperati
FROM corsi join pianidistudio p on codice=p.corso
       join esami e on codice=e.corso
GROUP BY codice, titolo, CFU;
```

4. Formulare in SQL l'interrogazione che restituisce, per ciascun corso, il codice, il titolo, il numero di studenti che lo hanno in piano di studi e il numero di studenti che hanno superato l'esame (cioè ciò che probabilmente si desiderava ottenere con l'interrogazione precedente)

Risultato desiderato:

	codice integer	titolo character (20)	cfu integer	numeroesamiprevisti bigint	numeroesamisuperati bigint
1	101	Fondamenti	12	4	3
2	102	POO	9	4	3
3	103	Basi di dati	6	2	1
4	104	Mobile Comp	6	2	1

Domanda 4 (20%) Lo schema concettuale seguente rappresenta un insieme di corsi e un insieme di iscrizioni a tali corsi. Nello schema l'attributo *Ricavo* è ridondante perché può essere ottenuto moltiplicando il costo del corso per il numero di iscritti (cioè il prodotto del valore dell'attributo *Costo* di ogni occorrenza dell'entità *Corso* per il numero di occorrenze dell'entità *Persona* a cui essa è correlata tramite la relationship *Iscrizione*).



Valutare se convenga o meno mantenere la ridondanza, assumendo che *Corso* abbia $N_C = 1.000$ occorrenze, *Persona* ne abbia $N_P = 3.000$ e *Iscrizione* $N_I = 10.000$ (e quindi che ogni corso abbia mediamente $N_{CI} = 10$ iscritti) e che il carico applicativo includa come operazioni principali le seguenti:

1. calcolo del ricavo di un corso (dato il codice), con frequenza $f_1 = 1.000$
2. iscrizione di una persona (già nella base di dati e di cui è dato il codice fiscale) ad un corso (di cui è dato il codice), con frequenza $f_2 = 100$; in sostanza, questa operazione richiede l'inserimento di una occorrenza di *Iscrizione*

Considerare i costi delle letture e scritture delle occorrenze di entità e relationship e considerare il costo delle scritture pari a quello delle letture.

Basi di dati — 25 gennaio 2018 — Compito C

Tempo a disposizione: un'ora per la prova breve, due ore per la prova lunga

Cognome: _____ Nome: _____ Matricola: _____

Basi di dati I — 25 gennaio 2018 — Compito C

Domanda 1 (50% per le prova breve e 25% per la prova completa)

Considerare la relazione seguente

Cod	Titolo	CC	Città	CN	Nazione	CM	Materia	CA	Area
1	Articolazioni	C1	Milano	N1	Italia	M1	Ortopedia	A1	Medicina
2	Fratture	C2	Roma	N1	Italia	M1	Ortopedia	A1	Medicina
3	Cuore2018	C1	Milano	N1	Italia	M2	Cardiologia	A1	Medicina
4	DB2018	C3	Monaco	N2	Germania	M1	Basi di dati	A2	Informatica

La relazione mostra (in forma non normalizzata) i dati di interesse in un certo anno per un insieme di congressi scientifici, secondo le seguenti specifiche:

- ogni congresso ha un codice e un titolo e si tiene una città
- ogni città ha un nome, si trova in una nazione e ha un codice che la identifica univocamente
- ogni nazione ha un codice e un nome
- ogni congresso è relativo ad una materia specifica (ad esempio l'ortopedia) che a sua volta appartiene ad un'area scientifica (ad esempio la medicina); ogni materia ha un nome e un codice che la identifica nell'ambito dell'area scientifica
- ogni area scientifica ha un codice e un nome

Con riferimento alle specifiche e ai dati nell'esempio

1. mostrare uno schema concettuale per la realtà di interesse (rispettando le specifiche per i vari codici)

2. mostrare una buona decomposizione della relazione originaria che soddisfi la BCNF (mostrare le tabelle, indicando la chiave di ciascuna con la sottolineatura)

Domanda 2 (50% per le prova breve e 25% per la prova completa)

Mostrare uno schema concettuale ancora per una realtà relativa a congressi, ma con le specifiche seguenti, che estendono quelle della domanda precedente:

- interessano congressi svolti anche in anni diversi
- ogni congresso appartiene ad una serie, e per ogni serie (ad esempio, la serie VLDB) esiste al più un congresso all'anno (quindi ad esempio esiste il VLDB 2017); ogni serie ha un codice e un titolo ed è relativa ad una materia (che ha codice, nome e area scientifica con gli stessi dettagli della domanda precedente)
- ogni congresso si svolge in una città (che ha codice, nome e nazione con gli stessi dettagli della domanda precedente), in una certa data (per semplicità un solo giorno)
- ogni congresso è composto di una serie di sessioni, ognuna con codice (univoco nella base di dati), orario e un insieme di partecipanti, ognuno dei quali presenta una relazione con un titolo (semplice stringa di testo) e ha una affiliazione (cioè una istituzione, ad esempio una università o una azienda); si noti che l'affiliazione di una stessa persona può variare da un congresso all'altro
- ogni istituzione ha un codice un nome e una città in cui ha sede

Domanda 3 (30%) Considerare la seguente porzione di base di dati che contiene alcune informazioni su corsi universitari, esami e piani di studio (ovviamente tanto le relazioni quanto gli attributi e anche le ennuple sono solo una parte di quelli di interesse; in particolare, la relazione piani di studio contiene le coppie studente-corso per le quali lo studente ha il corso nel proprio piano di studio)

CORSI		
<u>Codice</u>	Titolo	CFU
101	Fondamenti	12
102	POO	9
103	Basi di dati	6
104	Mobile Comp	6

ESAMI		
<u>Corso</u>	<u>Studente</u>	Voto
101	100001	30
101	100002	28
101	100003	26
102	100001	30
102	100002	27
102	100003	30
103	100001	30
104	100003	28

PIANIDISTUDIO	
<u>Corso</u>	<u>Studente</u>
101	100001
101	100002
101	100003
101	100004
102	100001
102	100002
102	100003
102	100004
103	100001
103	100002
104	100003
104	100004

1. Mostrare il risultato della seguente interrogazione

```
SELECT corso, titolo, CFU, count(studente) AS numeroesamiprevisti
FROM corsi join pianidistudio on codice=corso
GROUP BY corso, titolo, CFU;
```

2. Formulare in SQL l'interrogazione che conta, per ogni corso, quanti studenti lo hanno in piano di studio e non hanno ancora superato l'esame

3. Mostrare il risultato della seguente interrogazione

```
SELECT codice, titolo, CFU,
       count(e.studente) AS numeroesamiprevisti, count(e.studente) AS numeroesamisuperati
FROM corsi join pianidistudio p on codice=p.corso
       join esami e on codice=e.corso
GROUP BY codice, titolo, CFU;
```

4. Formulare in SQL l'interrogazione che restituisce, per ciascun corso, il codice, il titolo, il numero di studenti che lo hanno in piano di studi e il numero di studenti che hanno superato l'esame (cioè ciò che probabilmente si desiderava ottenere con l'interrogazione precedente)

Risultato desiderato:

	codice integer	titolo character (20)	cfu integer	numeroesamiprevisti bigint	numeroesamisuperati bigint
1	101	Fondamenti	12	4	3
2	102	POO	9	4	3
3	103	Basi di dati	6	2	1
4	104	Mobile Comp	6	2	1

Domanda 4 (20%) Lo schema concettuale seguente rappresenta un insieme di corsi e un insieme di iscrizioni a tali corsi. Nello schema l'attributo *Ricavo* è ridondante perché può essere ottenuto moltiplicando il costo del corso per il numero di iscritti (cioè il prodotto del valore dell'attributo *Costo* di ogni occorrenza dell'entità *Corso* per il numero di occorrenze dell'entità *Persona* a cui essa è correlata tramite la relationship *Iscrizione*).



Valutare se convenga o meno mantenere la ridondanza, assumendo che *Corso* abbia $C_C = 1.000$ occorrenze, *Persona* ne abbia $C_P = 3.000$ e *Iscrizione* $C_I = 20.000$ (e quindi che ogni corso abbia mediamente $C_{CI} = 20$ iscritti) e che il carico applicativo includa come operazioni principali le seguenti:

1. calcolo del ricavo di un corso (dato il codice), con frequenza $f_1 = 25$
2. iscrizione di una persona (già nella base di dati e di cui è dato il codice fiscale) ad un corso (di cui è dato il codice), con frequenza $f_2 = 10.000$; in sostanza, questa operazione richiede l'inserimento di una occorrenza di *Iscrizione*

Considerare i costi delle letture e scritture delle occorrenze di entità e relationship e considerare il costo delle scritture pari a quello delle letture.

Basi di dati — 25 gennaio 2018 — Compito D

Tempo a disposizione: un'ora per la prova breve, due ore per la prova lunga

Cognome: _____ Nome: _____ Matricola: _____

Basi di dati I — 25 gennaio 2018 — Compito D

Domanda 1 (50% per le prova breve e 25% per la prova completa)

Considerare la relazione seguente

Cod	Titolo	CC	Città	CN	Nazione	CM	Materia	CA	Area
1	Articolazioni	C1	Milano	N1	Italia	M1	Ortopedia	A1	Medicina
2	Fratture	C2	Roma	N1	Italia	M1	Ortopedia	A1	Medicina
3	Cuore2018	C1	Milano	N1	Italia	M2	Cardiologia	A1	Medicina
4	DB2018	C1	Monaco	N2	Germania	M3	Basi di dati	A2	Informatica

La relazione mostra (in forma non normalizzata) i dati di interesse in un certo anno per un insieme di convegni scientifici, secondo le seguenti specifiche:

- ogni convegno ha un codice e un titolo e si tiene una città
- ogni città ha un nome, si trova in una nazione e ha un codice che la identifica nell'ambito della nazione
- ogni nazione ha un codice e un nome
- ogni convegno è relativo ad una materia specifica (ad esempio l'ortopedia) che a sua volta appartiene ad un'area scientifica (ad esempio la medicina); ogni materia ha un nome e un codice che la identifica univocamente
- ogni area scientifica ha un codice e un nome

Con riferimento alle specifiche e ai dati nell'esempio

1. mostrare uno schema concettuale per la realtà di interesse (rispettando le specifiche per i vari codici)

2. mostrare una buona decomposizione della relazione originaria che soddisfi la BCNF (mostrare le tabelle, indicando la chiave di ciascuna con la sottolineatura)

Domanda 2 (50% per la prova breve e 25% per la prova completa)

Mostrare uno schema concettuale ancora per una realtà relativa a convegni, ma con le specifiche seguenti, che estendono quelle della domanda precedente:

- interessano convegni svolti anche in anni diversi
- ogni convegno appartiene ad una serie, e per ogni serie (ad esempio, la serie VLDB) esiste al più un convegno all'anno (quindi ad esempio esiste il VLDB 2017); ogni serie ha un codice e un titolo ed è relativa ad una materia (che ha codice, nome e area scientifica con gli stessi dettagli della domanda precedente)
- ogni convegno si svolge in una città (che ha codice, nome e nazione con gli stessi dettagli della domanda precedente), in una certa data (per semplicità un solo giorno)
- ogni convegno è composto di una serie di sessioni, ognuna con codice (univoco nella base di dati), orario e un insieme di partecipanti, ognuno dei quali presenta una relazione con un titolo (semplice stringa di testo) e ha una affiliazione (cioè una istituzione, ad esempio una università o una azienda); si noti che l'affiliazione di una stessa persona può variare da un convegno all'altro
- ogni istituzione ha un codice un nome e una città in cui ha sede

Domanda 3 (30%) Considerare la seguente porzione di base di dati che contiene alcune informazioni su corsi universitari, esami e piani di studio (ovviamente tanto le relazioni quanto gli attributi e anche le ennuple sono solo una parte di quelli di interesse; in particolare, la relazione piani di studio contiene le coppie studente-corso per le quali lo studente ha il corso nel proprio piano di studio)

CORSI		
<u>Codice</u>	Titolo	CFU
101	Fondamenti	12
102	POO	9
103	Basi di dati	6
104	Mobile Comp	6

ESAMI		
<u>Corso</u>	<u>Studente</u>	Voto
101	100001	30
101	100002	28
101	100003	26
102	100001	30
102	100002	27
102	100003	30
103	100001	30
104	100003	28

PIANIDISTUDIO	
<u>Corso</u>	<u>Studente</u>
101	100001
101	100002
101	100003
101	100004
102	100001
102	100002
102	100003
102	100004
103	100001
103	100002
104	100003
104	100004

1. Mostrare il risultato della seguente interrogazione

```
SELECT corso, titolo, CFU, count(studente) AS numeroesamiprevisti
FROM corsi join pianidistudio on codice=corso
GROUP BY corso, titolo, CFU;
```

2. Formulare in SQL l'interrogazione che conta, per ogni corso, quanti studenti lo hanno in piano di studio e non hanno ancora superato l'esame

3. Mostrare il risultato della seguente interrogazione

```
SELECT codice, titolo, CFU,
       count(e.studente) AS numeroesamiprevisti, count(e.studente) AS numeroesamisuperati
FROM corsi join pianidistudio p on codice=p.corso
       join esami e on codice=e.corso
GROUP BY codice, titolo, CFU;
```

4. Formulare in SQL l'interrogazione che restituisce, per ciascun corso, il codice, il titolo, il numero di studenti che lo hanno in piano di studi e il numero di studenti che hanno superato l'esame (cioè ciò che probabilmente si desiderava ottenere con l'interrogazione precedente)

Risultato desiderato:

	codice integer	titolo character (20)	cfu integer	numeroesamiprevisti bigint	numeroesamisuperati bigint
1	101	Fondamenti	12	4	3
2	102	POO	9	4	3
3	103	Basi di dati	6	2	1
4	104	Mobile Comp	6	2	1

Domanda 4 (20%) Lo schema concettuale seguente rappresenta un insieme di corsi e un insieme di iscrizioni a tali corsi. Nello schema l'attributo *Ricavo* è ridondante perché può essere ottenuto moltiplicando il costo del corso per il numero di iscritti (cioè il prodotto del valore dell'attributo *Costo* di ogni occorrenza dell'entità *Corso* per il numero di occorrenze dell'entità *Persona* a cui essa è correlata tramite la relationship *Iscrizione*).



Valutare se convenga o meno mantenere la ridondanza, assumendo che *Corso* abbia $N_C = 1.000$ occorrenze, *Persona* ne abbia $N_P = 3.000$ e *Iscrizione* $N_I = 10.000$ (e quindi che ogni corso abbia mediamente $N_{CI} = 10$ iscritti) e che il carico applicativo includa come operazioni principali le seguenti:

1. calcolo del ricavo di un corso (dato il codice), con frequenza $f_1 = 2.000$
2. iscrizione di una persona (già nella base di dati e di cui è dato il codice fiscale) ad un corso (di cui è dato il codice), con frequenza $f_2 = 200$; in sostanza, questa operazione richiede l'inserimento di una occorrenza di *Iscrizione*

Considerare i costi delle letture e scritture delle occorrenze di entità e relationship e considerare il costo delle scritture pari a quello delle letture.

Basi di dati — 25 gennaio 2018 — Compito A

Possibili soluzioni

Tempo a disposizione: un'ora per la prova breve, due ore per la prova lunga

Cognome: _____ **Nome:** _____ **Matricola:** _____

Domanda 1 (50% per le prova breve e 25% per la prova completa)

Considerare la relazione seguente

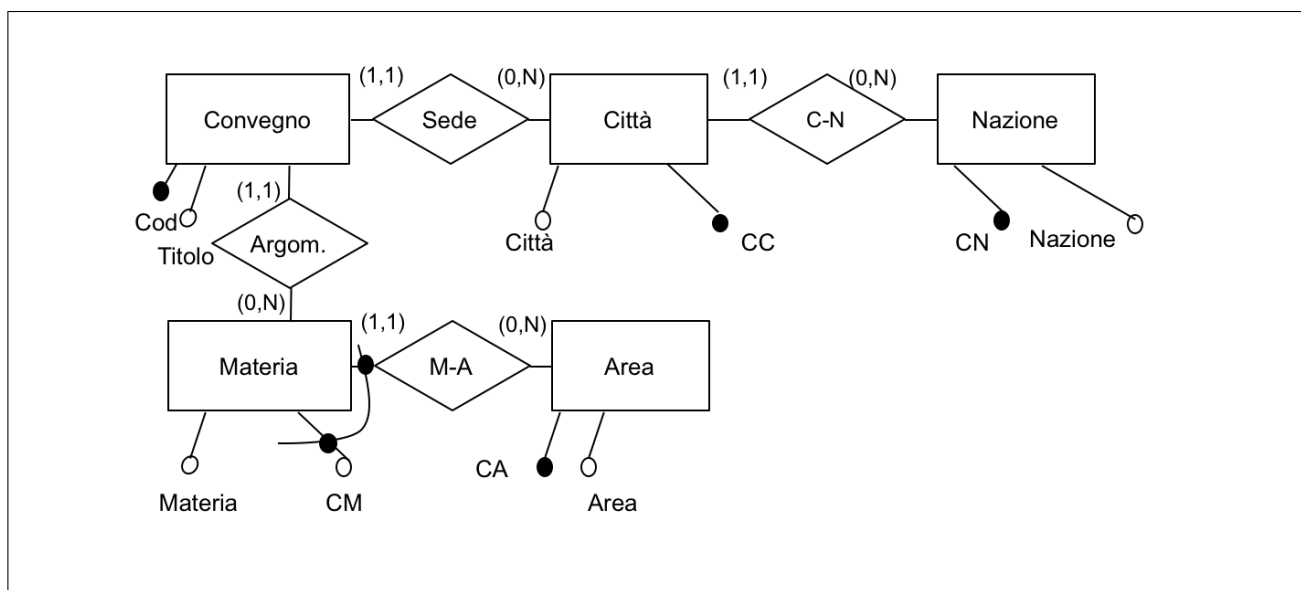
Cod	Titolo	CC	Città	CN	Nazione	CM	Materia	CA	Area
1	Articolazioni	C1	Milano	N1	Italia	M1	Ortopedia	A1	Medicina
2	Fratture	C2	Roma	N1	Italia	M1	Ortopedia	A1	Medicina
3	Cuore2018	C1	Milano	N1	Italia	M2	Cardiologia	A1	Medicina
4	DB2018	C3	Monaco	N2	Germania	M1	Basi di dati	A2	Informatica

La relazione mostra (in forma non normalizzata) i dati di interesse in un certo anno per un insieme di congressi scientifici, secondo le seguenti specifiche:

- ogni congresso ha un codice e un titolo e si tiene una città
- ogni città ha un nome, si trova in una nazione e ha un codice che la identifica univocamente
- ogni nazione ha un codice e un nome
- ogni congresso è relativo ad una materia specifica (ad esempio l'ortopedia) che a sua volta appartiene ad un'area scientifica (ad esempio la medicina); ogni materia ha un nome e un codice che la identifica nell'ambito dell'area scientifica
- ogni area scientifica ha un codice e un nome

Con riferimento alle specifiche e ai dati nell'esempio

1. mostrare uno schema concettuale per la realtà di interesse (rispettando le specifiche per i vari codici)



2. mostrare una buona decomposizione della relazione originaria che soddisfi la BCNF (mostrare le tabelle, indicando la chiave di ciascuna con la sottolineatura)

CONVEGNI				
<u>Cod</u>	Titolo	CC	CM	CA
1	Articolazioni	C1	M1	A1
2	Fratture	C2	M1	A1
3	Cuore2018	C1	M2	A1
4	DB2018	C3	M1	A2

CITTÀ		
<u>CC</u>	Città	CN
C1	Milano	N1
C2	Roma	N1
C3	Monaco	N2

MATERIE		
<u>CM</u>	Materia	<u>CA</u>
M1	Ortopedia	A1
M2	Cardiologia	A1
M1	Basi di dati	A2

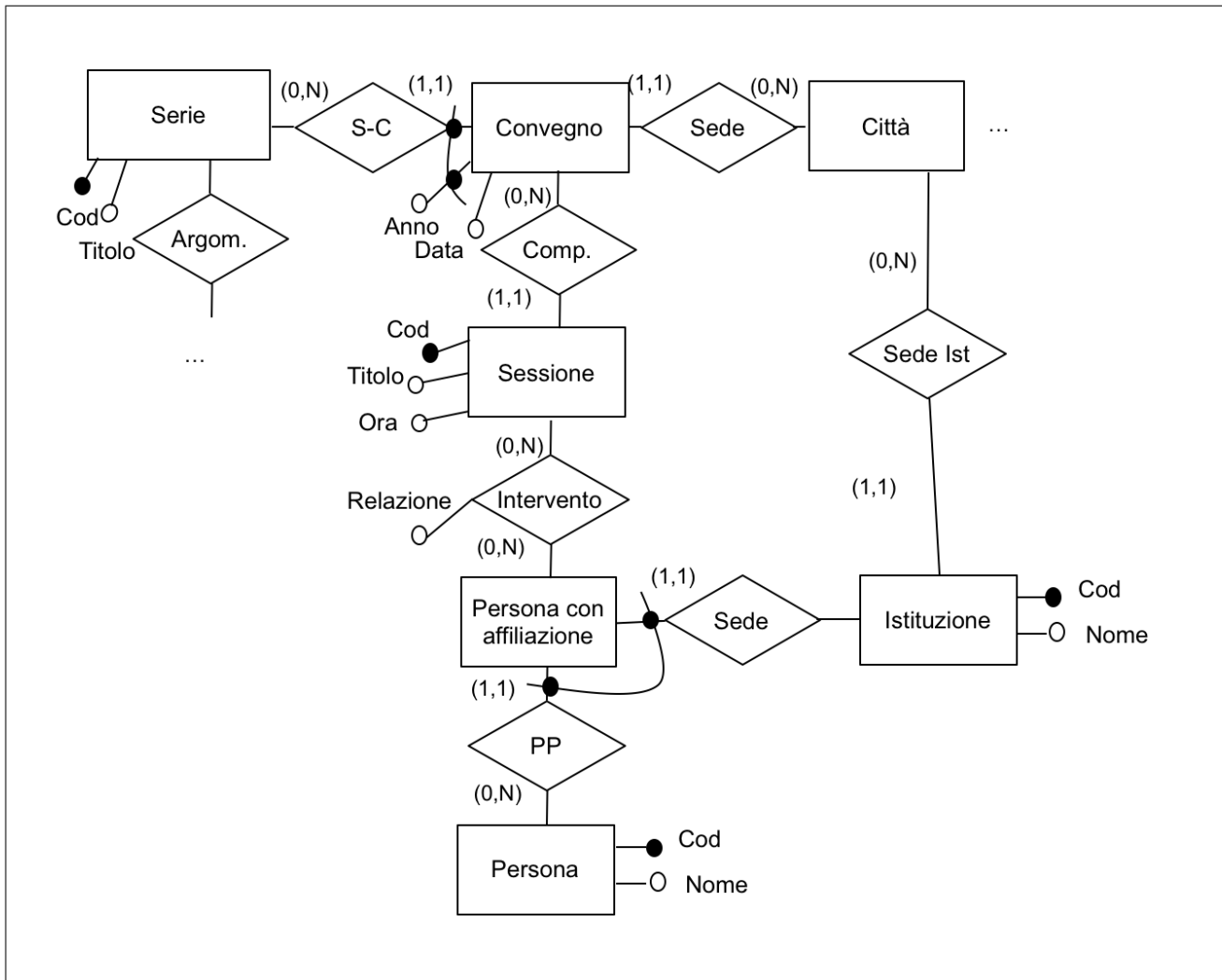
NAZIONI	
<u>CN</u>	Nazione
N1	Italia
N2	Germania

AREE	
<u>CA</u>	Area
A1	Medicina
A2	Informatica

Domanda 2 (50% per le prova breve e 25% per la prova completa)

Mostrare uno schema concettuale ancora per una realtà relativa a congressi, ma con le specifiche seguenti, che estendono quelle della domanda precedente:

- interessano congressi svolti anche in anni diversi
- ogni congresso appartiene ad una serie, e per ogni serie (ad esempio, la serie VLDB) esiste al più un congresso all'anno (quindi ad esempio esiste il VLDB 2017); ogni serie ha un codice e un titolo ed è relativa ad una materia (che ha codice, nome e area scientifica con gli stessi dettagli della domanda precedente)
- ogni congresso si svolge in una città (che ha codice, nome e nazione con gli stessi dettagli della domanda precedente), in una certa data (per semplicità un solo giorno)
- ogni congresso è composto di una serie di sessioni, ognuna con codice (univoco nella base di dati), orario e un insieme di partecipanti, ognuno dei quali presenta una relazione con un titolo (semplice stringa di testo) e ha una affiliazione (cioè una istituzione, ad esempio una università o una azienda); si noti che l'affiliazione di una stessa persona può variare da un congresso all'altro
- ogni istituzione ha un codice un nome e una città in cui ha sede



Basi di dati I — 25 gennaio 2018 — Compito A

Domanda 3 (30%) Considerare la seguente porzione di base di dati che contiene alcune informazioni su corsi universitari, esami e piani di studio (ovviamente tanto le relazioni quanto gli attributi e anche le ennuple sono solo una parte di quelli di interesse; in particolare, la relazione piani di studio contiene le coppie studente-corso per le quali lo studente ha il corso nel proprio piano di studio)

CORSI			ESAMI			PIANIDISTUDIO	
Codice	Titolo	CFU	Corso	Studente	Voto	Corso	Studente
101	Fondamenti	12	101	100001	30	101	100001
102	POO	9	101	100002	28	101	100002
103	Basi di dati	6	101	100003	26	101	100003
104	Mobile Comp	6	102	100001	30	101	100004
			102	100002	27	102	100001
			102	100003	30	102	100002
			103	100001	30	102	100003
			104	100003	28	102	100004
						103	100001
						103	100002
						104	100003
						104	100004

1. Mostrare il risultato della seguente interrogazione

```
SELECT corso, titolo, CFU, count(studente) AS numeroesamiprevisti
FROM corsi join pianidistudio on codice=corso
GROUP BY corso, titolo, CFU;
```

	corso integer	titolo character (20)	cfu integer	numeroesamiprevisti bigint
1	101	Fondamenti	12	4
2	102	POO	9	4
3	103	Basi di dati	6	2
4	104	Mobile Comp	6	2

2. Formulare in SQL l'interrogazione che conta, per ogni corso, quanti studenti lo hanno in piano di studio e non hanno ancora superato l'esame

```
CREATE VIEW esamiNonSuperati AS
SELECT corso, studente
FROM pianidistudio
EXCEPT
SELECT corso, studente
FROM esami

SELECT codice, titolo, count(studente)
FROM corsi join esamiNonSuperati on codice=corso
GROUP BY codice, titolo
```

3. Mostrare il risultato della seguente interrogazione

```
SELECT codice, titolo, CFU,
       count(e.studente) AS numeroesamiprevisti, count(e.studente) AS numeroesamisuperati
FROM corsi join pianidistudio p on codice=p.corso
       join esami e on codice=e.corso
GROUP BY codice, titolo, CFU;
```

	codice integer	titolo character (20)	cfu integer	numeroesamiprevisti bigint	numeroesamisuperati bigint
1	101	Fondamenti	12	12	12
2	102	POO	9	12	12
3	103	Basi di dati	6	2	2
4	104	Mobile Comp	6	2	2

4. Formulare in SQL l'interrogazione che restituisce, per ciascun corso, il codice, il titolo, il numero di studenti che lo hanno in piano di studi e il numero di studenti che hanno superato l'esame (cioè ciò che probabilmente si desiderava ottenere con l'interrogazione precedente)

Risultato desiderato:

	codice integer	titolo character (20)	cfu integer	numeroesamiprevisti bigint	numeroesamisuperati bigint
1	101	Fondamenti	12	4	3
2	102	POO	9	4	3
3	103	Basi di dati	6	2	1
4	104	Mobile Comp	6	2	1

```
CREATE OR REPLACE VIEW esamiprevisti AS
SELECT pianidistudio.corso,
       count(pianidistudio.studente) AS numeroesamiprevisti
FROM pianidistudio
GROUP BY pianidistudio.corso;
```

```
CREATE OR REPLACE VIEW esamisuperati AS
SELECT esami.corso,
       count(esami.studente) AS numeroesamisuperati
FROM esami
GROUP BY esami.corso;
```

```
SELECT codice, titolo, CFU, numeroesamiprevisti, numeroesamisuperati
FROM corsi join esamiprevisti p on codice=p.corso
       join esamisuperati e on codice=e.corso
```

Domanda 4 (20%) Lo schema concettuale seguente rappresenta un insieme di corsi e un insieme di iscrizioni a tali corsi. Nello schema l'attributo *Ricavo* è ridondante perché può essere ottenuto moltiplicando il costo del corso per il numero di iscritti (cioè il prodotto del valore dell'attributo *Costo* di ogni occorrenza dell'entità *Corso* per il numero di occorrenze dell'entità *Persona* a cui essa è correlata tramite la relationship *Iscrizione*).



Valutare se convenga o meno mantenere la ridondanza, assumendo che *Corso* abbia $C_C = 1.000$ occorrenze, *Persona* ne abbia $C_P = 3.000$ e *Iscrizione* $C_I = 30.000$ (e quindi che ogni corso abbia mediamente $C_{CI} = 30$ iscritti) e che il carico applicativo includa come operazioni principali le seguenti:

1. calcolo del ricavo di un corso (dato il codice), con frequenza $f_1 = 50$
2. iscrizione di una persona (già nella base di dati e di cui è dato il codice fiscale) ad un corso (di cui è dato il codice), con frequenza $f_2 = 20.000$; in sostanza, questa operazione richiede l'inserimento di una occorrenza di *Iscrizione*

Considerare i costi delle letture e scritture delle occorrenze di entità e relationship e considerare il costo delle scritture pari a quello delle letture.

Soluzione

Dati numerici per il compito A

con ridondanza $c_1 \times f_1 + c_2 \times f_2 = 1 \times 50 + 3 \times 20.000 = \text{ca. } 60.000$

- c_1 è pari a 1 perché si deve leggere un'occorrenza di *Corso*
- c_2 è pari a 3 perché si deve inserire un'occorrenza di *Iscrizione* e si deve leggere e scrivere un'occorrenza di *Corso*

senza ridondanza $c_1 \times f_1 + c_2 \times f_2 = 30 \times 50 + 1 \times 20.000 = \text{ca. } 21.500$

- c_1 è pari a $C_{CI} = \text{ca. } 30$ perché si debbono leggere C_{CI} occorrenze di *Iscrizione*
- c_2 è pari a 1 perché si deve inserire un'occorrenza di *Iscrizione*

Quindi risulta conveniente non introdurre la ridondanza

Per il compito C, soluzione molto simile, con frequenze dimezzate. Per i compiti B e D, essendo scambiate le frequenze, ma analoghe le operazioni, la ridondanza conviene

Basi di dati — 25 gennaio 2018 — Compito B

Possibili soluzioni

Tempo a disposizione: un'ora per la prova breve, due ore per la prova lunga

Cognome: _____ **Nome:** _____ **Matricola:** _____

Domanda 1 (50% per le prova breve e 25% per la prova completa)

Considerare la relazione seguente

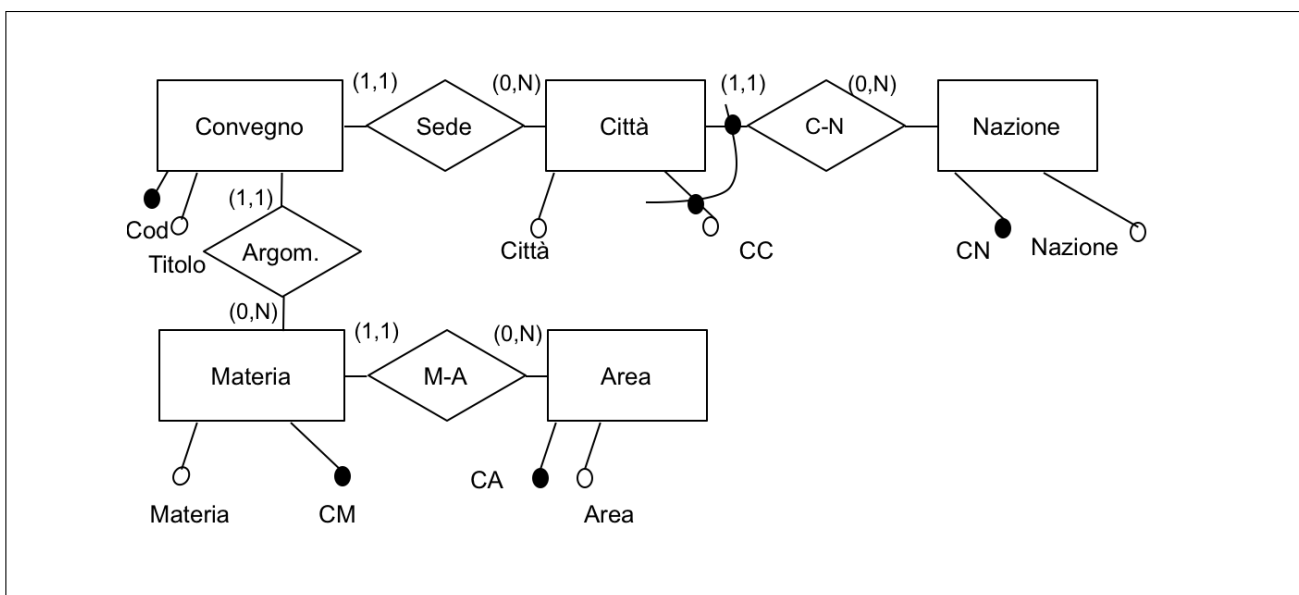
Cod	Titolo	CC	Città	CN	Nazione	CM	Materia	CA	Area
1	Articolazioni	C1	Milano	N1	Italia	M1	Ortopedia	A1	Medicina
2	Fratture	C2	Roma	N1	Italia	M1	Ortopedia	A1	Medicina
3	Cuore2018	C1	Milano	N1	Italia	M2	Cardiologia	A1	Medicina
4	DB2018	C1	Monaco	N2	Germania	M3	Basi di dati	A2	Informatica

La relazione mostra (in forma non normalizzata) i dati di interesse in un certo anno per un insieme di convegni scientifici, secondo le seguenti specifiche:

- ogni convegno ha un codice e un titolo e si tiene una città
- ogni città ha un nome, si trova in una nazione e ha un codice che la identifica nell'ambito della nazione
- ogni nazione ha un codice e un nome
- ogni convegno è relativo ad una materia specifica (ad esempio l'ortopedia) che a sua volta appartiene ad un'area scientifica (ad esempio la medicina); ogni materia ha un nome e un codice che la identifica univocamente
- ogni area scientifica ha un codice e un nome

Con riferimento alle specifiche e ai dati nell'esempio

1. mostrare uno schema concettuale per la realtà di interesse (rispettando le specifiche per i vari codici)



2. mostrare una buona decomposizione della relazione originaria che soddisfi la BCNF (mostrare le tabelle, indicando la chiave di ciascuna con la sottolineatura)

CONVEGNI				
<u>Cod</u>	Titolo	CC	CN	CM
1	Articolazioni	C1	N1	M1
2	Fratture	C2	N1	M1
3	Cuore2018	C1	N1	M2
4	VLDB	C1	N1	M3

CITTÀ		
<u>CC</u>	Città	CN
C1	Milano	N1
C2	Roma	N1
C1	Monaco	N2

MATERIE		
<u>CM</u>	Materia	CA
M1	Ortopedia	A1
M2	Cardiologia	A1
M3	Basi di dati	A2

NAZIONI	
<u>CN</u>	Nazione
N1	Italia
N2	Germania

AREE	
<u>CA</u>	Area
A1	Medicina
A2	Informatica

Domanda 2 (50% per le prova breve e 25% per la prova completa)

Mostrare uno schema concettuale ancora per una realtà relativa a convegni, ma con le specifiche seguenti, che estendono quelle della domanda precedente:

- interessano convegni svolti anche in anni diversi
- ogni convegno appartiene ad una serie, e per ogni serie (ad esempio, la serie VLDB) esiste al più un convegno all'anno (quindi ad esempio esiste il VLDB 2017); ogni serie ha un codice e un titolo ed è relativa ad una materia (che ha codice, nome e area scientifica con gli stessi dettagli della domanda precedente)
- ogni convegno si svolge in una città (che ha codice, nome e nazione con gli stessi dettagli della domanda precedente), in una certa data (per semplicità un solo giorno)
- ogni convegno è composto di una serie di sessioni, ognuna con codice (univoco nella base di dati), orario e un insieme di partecipanti, ognuno dei quali presenta una relazione con un titolo (semplice stringa di testo) e ha una affiliazione (cioè una istituzione, ad esempio una università o una azienda); si noti che l'affiliazione di una stessa persona può variare da un convegno all'altro
- ogni istituzione ha un codice un nome e una città in cui ha sede

