

## Basi di dati II, primo modulo Prova parziale — 22 aprile 2010 — Compito A Cenni sulle soluzioni

Rispondere su questo fascicolo. Consegnare anche i fogli usati come brutta copia, che però non verranno corretti.  
Tempo a disposizione: un'ora e trenta minuti.

Cognome \_\_\_\_\_ Nome \_\_\_\_\_ Matricola \_\_\_\_\_ Ordin. \_\_\_\_\_

**Domanda 1** Considerare un disco con una velocità di rotazione di  $v = 10.000$  giri al minuto e un tempo medio di posizionamento della testina (tempo di seek)  $t_S = 5$  msec. Ogni traccia contiene  $N = 240$  blocchi e ogni blocco contiene  $B = 4$  KB. Rispondere alle seguenti domande mostrando formula e valore numerico numero (N.B. non servono calcolatrici, i risultati sono semplici, approssimare  $1 \text{ MB} = 1000 \text{ KB}$ ).

1. Qual è tempo medio di latenza (attesa dovuta alla rotazione)  $t_L$ ?  
è il tempo necessario per mezzo giro:  $1/2 \times 1/v \times 60 = 1/2 \times 1/10.000 \times 60 = 0,003 \text{ sec} = 3 \text{ msec}$
2. Qual è la massima velocità di trasferimento (in MB al secondo)?  
 $v/60 \times N \times B = 10.000/60 \times 240 \times 4 \text{ KB} = 160.000 \text{ KB/sec} = 160 \text{ MB/sec}$
3. Qual è il tempo minimo  $t_B$  necessario per leggere un blocco?  
 $1/v \times 60 \times 1/N = 1/10.000 \times 60 \times 1/240 = 1/10.000 \times 1/240 \times 60 \text{ sec} = 25 \mu\text{sec}$
4. Qual è il tempo medio necessario per leggere un blocco?  
 $t_S + t_L + t_B = 5 + 3 + 0,025 \text{ msec} = \text{ca } 8 \text{ msec}$
5. Qual è il tempo medio necessario per leggere  $n = 40$  blocchi contigui?  
 $t_S + t_L + n \times t_B = 5 + 3 + 40 \times 0,025 \text{ msec} = \text{ca } 9 \text{ msec}$

**Domanda 2** Una catena di supermercati ha una base di dati dei propri clienti che dispongono di una “tessera fedeltà,” con varie informazioni su ciascun cliente, fra cui (a) il totale dei punti acquisiti attraverso l’uso della tessera e (b) il negozio della catena cui fa riferimento (ad esempio, quello presso cui ha inizialmente richiesto la tessera). Si vuole eseguire su di essa l’interrogazione che calcola, per ciascun negozio, il numero dei clienti, la somma dei punti fedeltà dei clienti e la relativa media per cliente. Indicare quale livello di isolamento (READ UNCOMMITTED, READ COMMITTED, REPEATABLE READ o SERIALIZABLE) si potrebbe scegliere in ciascuno dei seguenti casi (si supponga che, in generale, sia stato rilevato che, nel corso degli inserimenti e delle modifiche, vengono inseriti valori sbagliati anche di vari ordini di grandezza, che sono poi corretti prima del commit):

1. L’operazione è eseguita mentre vengono inseriti alcuni nuovi clienti (per ciascun negozio pochi rispetto a quelli già presenti), con la finalità di acquisire informazioni approssimate ma ragionevolmente indicative sugli andamenti complessivi.
2. L’operazione è eseguita mentre vengono inseriti molti nuovi clienti, con la finalità di acquisire informazioni approssimate ma ragionevolmente indicative sugli andamenti complessivi.
3. L’operazione è eseguita mentre vengono inseriti alcuni nuovi clienti (per ciascun negozio pochi rispetto a quelli già presenti), con la finalità di individuare i primi tre negozi da premiare in una campagna promozionale sulla base dei punti acquisiti dai rispettivi clienti.
4. L’operazione è eseguita mentre vengono modificati i valori dei punti fedeltà di tutti i clienti (a seguito di una ridefinizione dei criteri di assegnazione dei punti stessi), con la finalità di individuare i primi tre negozi da premiare in una campagna promozionale sulla base dei punti acquisiti dai rispettivi clienti.
5. L’operazione è eseguita in un momento in cui non ci sono aggiornamenti di alcun genere, con la finalità di acquisire informazioni approssimate ma ragionevolmente indicative sugli andamenti complessivi.

Risposte				
1.	2.	3.	4.	5.
RC	S	S	RR	RU

**Domanda 3** Molti DBMS utilizzano, oltre a quelli basati sui B<sup>+</sup>-tree, anche indici basati su una tecnica detta “bitmap.” Un indice bitmap, su un attributo A con a valori diversi, per una relazione con N<sub>R</sub> ennuple, è costituito da a vettori (uno per ciascun valore di A) di N<sub>R</sub> bit ciascuno (un bit per ogni ennupla della relazione): l’i-esimo bit del vettore associato al valore a<sub>j</sub> è 1 se il valore della i-esima ennupla sull’attributo A è a<sub>j</sub> e 0 se il valore è diverso. L’accesso ai vettori è organizzato per mezzo di un albero. In sostanza, rispetto ad un B<sup>+</sup>-tree, abbiamo una struttura di accesso simile, ma con foglie diverse: nel B<sup>+</sup>-tree, per ogni valore dell’attributo abbiamo una lista di indirizzi, mentre nell’indice bitmap abbiamo un vettore di bit. Esiste poi una struttura (tabella, indice o altro) che associa gli indirizzi delle ennuple ai numeri da 1 a N<sub>R</sub>. Sulla base degli elementi forniti, con riferimento ad un sistema con blocchi di dimensione B = 4 KB e indirizzi dei record di lunghezza b = 8 B e ad una relazione R con N<sub>R</sub> = 1.600.000 ennuple e (fra gli altri) quattro attributi, A, C, E, F ciascuno con a = 20 valori diversi e una lunghezza c = 6 byte

1. valutare lo spazio SBM (in blocchi) necessario per le foglie di un indice bitmap su A; valutare anche lo spazio S<sub>I</sub> necessario per una tabella (una sorta di array in memoria secondaria, ad esempio) che permetta di associare un indirizzo a ciascun valore compreso fra 1 e N<sub>R</sub>; per confronto, calcolare lo spazio ST necessario per le foglie di un B<sup>+</sup>-tree (con riferimenti ai record) per un indice su A

Poi, valutare (sia nel caso di indici bitmap, sia nel caso di B<sup>+</sup>-tree) il costo (in termini di numero di accessi a memoria secondaria, supponendo che tutti i nodi intermedi di tutti gli indici siano nei buffer; basta quindi contare il numero di foglie da visitare e l’accesso alle ennuple, quando necessario) per le seguenti interrogazioni

2. SELECT COUNT(\*) FROM R WHERE A = 5
3. SELECT \* FROM R WHERE A = 5
4. SELECT \* FROM R WHERE A = 15 AND C = 14 AND E = 13 AND F = 9 (considerare i valori indipendenti l’uno dall’altro; quindi assumere che il risultato contenga circa N<sub>R</sub>/a<sup>4</sup> ennuple)

Riportare le risposte nella tabella sottostante, nelle righe da 1 a 4, indicando formula e valore numerico.

Risposte	
	Indici B <sup>+</sup> -tree
1	$SBM = \frac{a \times N_R}{8 \times B} = \frac{20 \times 1.600.000}{8 \times 4000} = 1000$ $S_I = \frac{N_R \times b}{B} = \frac{1.600.000 \times 8}{4000} = 3200$
2	$ST = \frac{a \times c + N_R \times b}{B \times f} =$ $= \frac{20 \times 6 + 1.600.000 \times 8}{4000 \times 0,66} \approx 4800$ <p>(f è il fattore di riempimento del B<sup>+</sup>-tree assunto pari a 0,66)</p>
3	<p>Basta accedere ad un vettore:</p> $\frac{N_R}{8 \times B} = \frac{1.600.000}{8 \times 4000} = 50$ <p>Come sopra, più l’accesso ai record (attraverso la struttura intermedia)</p> $\frac{N_R}{8 \times B} + 2 \frac{N_R}{a} = 50 + 2 \times \frac{1.600.000}{20} \approx 160.000$ <p>Come sopra, più l’accesso ai record</p> $\frac{1}{a} \frac{N_R \times b}{B \times f} + \frac{N_R}{a} = \dots + \frac{1.600.000}{20} = 80.000$
4	<p>(Si suppone un indice per ciascuno degli attributi)</p> <p>Quattro volte il caso 2, più l’accesso ai record</p> $4 \times \frac{N_R}{8 \times B} + 2 \times \frac{N_R}{a^4} = 200 + 20$ <p>(Si suppone un indice per ciascuno degli attributi)</p> <p>Quattro volte il caso 2, più l’accesso ai record</p> $4 \times \frac{1}{a} \frac{N_R \times b}{B \times f} + \frac{N_R}{a^4} = 4 \times 240 + 20$

**Domanda 4** Specificare per ciascuno dei seguenti schedule, la sua appartenenza alle varie classi (scrivere “sì” o “no” nelle caselle della tabella; assumere che gli identificatori delle transazioni corrispondano ai timestamp).

Risposte					
	Seriale	VSR	CSR	TS	2PL
$r_1(y)r_2(z)r_2(y)w_2(y)w_2(z)r_1(z)$	no	no	no	no	no
$r_1(x)w_1(x)r_2(x)w_2(x)r_0(y)w_1(y)$	no	VSR	CSR	TS	no
$r_1(x)r_2(z)w_2(z)w_1(x)r_2(x)w_2(x)$	no	VSR	CSR	TS	2PL
$r_2(x)w_2(x)r_1(x)w_1(x)$	seriale	VSR	CSR	no	2PL

**Domanda 5** Considerare uno schema dimensionale relativo agli esami, che utilizzi, come tabella dei fatti e come una delle dimensioni, le relazioni come quelle qui schematizzate:

KStudente	KCorso	KData	Voto	...
301	201	405	25	...
301	202	406	28	...
302	201	407	30	...
302	203	407	22	...
...	...	...	...	...

KCorso	Titolo	Crediti	...
201	Fisica I	6	...
202	Chimica	9	...
203	Geometria	6	...
...	...	...	...

Supporre che si presentino le seguenti esigenze di modifica:

- i corsi cambiano nome nel tempo: per esempio, il corso nella prima ennupla potrebbe ad un certo punto cambiare nome da “Fisica I” in “Meccanica”; interessano selezioni e aggregazioni relative agli esami tanto con riferimento al nome del corso (al momento dell’esame) quanto alla sua identità (un codice che viene introdotto allo scopo, ma non sempre viene utilizzato, perché alcuni analisti preferiscono fare riferimento al nome corrente del corso); le modifiche sono rare, ma è possibile che ci siano corsi con vari cambiamenti di nome;
- per ogni corso, interessa rappresentare anche il docente, per supportare analisi sugli esami svolti da ciascun docente; i docenti cambiano nel tempo e passano da un corso all’altro (e possono anche tenere più corsi nello stesso momento, ma ogni corso ha, in un certo giorno, un solo docente); è disponibile l’informazione relativa ai docenti dei corsi nel tempo (per tutto il periodo, anche passato, di interesse).

Mostrare nuove versioni delle due tabelle che permettano di soddisfare le esigenze sopra citate (mostrare anche i dati, con riferimento a quelli presenti negli esempi sopra, aggiungendo nuovi dati ragionevoli, che permettano di comprendere le modifiche).

*Possibile soluzione*

La tabella dei fatti può essere estesa aggiungendo una dimensione supplementare, che dipende da Corso e Data e può quindi essere facilmente aggiunta anche ad una tabella dei fatti esistente (i dati, come detto, sono disponibili):

KStudente	KCorso	KData	KDocente	Voto	...
301	201	405	701	25	...
301	202	406	701	28	...
302	201	407	702	30	...
302	203	407	703	22	...
...	...	...	...	...	...

Per la dimensione può essere utile la tecnica della “slowly changing dimension,” con un nuovo elemento (quindi una ennupla nella relazione) per ogni modifica. Viste le specifiche, qui potrebbe essere utile introdurre un codice, che non cambia nel tempo, e avere due attributi per il nome, con quello attuale e quello “dell’epoca.”

KCorso	Codice	Titolo	TitoloAttuale	Crediti	...
201	FIS01	Fisica I	Meccanica	6	...
202	CHIM01	Chimica	Chimica	9	...
203	GEOM01	Geometria	Geometria	6	...
...	...	...	...	...	...
209	FIS01	Meccanica	Meccanica	6	...

## Basi di dati II, primo modulo Prova parziale — 22 aprile 2010 — Compito B Cenni sulle soluzioni

Rispondere su questo fascicolo. Consegnare anche i fogli usati come brutta copia, che però non verranno corretti.  
Tempo a disposizione: un'ora e trenta minuti.

Cognome \_\_\_\_\_ Nome \_\_\_\_\_ Matricola \_\_\_\_\_ Ordin. \_\_\_\_\_

**Domanda 1** Considerare un disco con una velocità di rotazione di  $r = 10.000$  giri al minuto e un tempo medio di posizionamento della testina (tempo di seek)  $t_S = 5$  msec. Ogni traccia contiene  $P = 300$  blocchi e ogni blocco contiene  $B = 4$  KB. Rispondere alle seguenti domande mostrando formula e valore numerico numero (N.B. non servono calcolatrici, i risultati sono semplici, approssimare  $1 \text{ MB} = 1000 \text{ KB}$ ).

1. Qual è tempo medio di latenza (attesa dovuta alla rotazione)  $t_L$ ?  
è il tempo necessario per mezzo giro:  $1/2 \times 1/r \times 60 = 1/2 \times 1/10.000 \times 60 = 0,003 \text{ sec} = 3 \text{ msec}$
2. Qual è la massima velocità di trasferimento (in MB al secondo)?  
 $r/60 \times P \times B = 10.000/60 \times 300 \times 4 \text{ KB} = 200.000 \text{ KB/sec} = 200 \text{ MB/sec}$
3. Qual è il tempo minimo  $t_B$  necessario per leggere un blocco?  
 $1/r \times 60 \times 1/P = 1/10.000 \times 60 \times 1/300 = 1/10.000 \times 1/300 \times 60 \text{ sec} = 20 \mu\text{sec}$
4. Qual è il tempo medio necessario per leggere un blocco?  
 $t_S + t_L + t_B = 5 + 3 + 0,020 \text{ msec} = \text{ca } 8 \text{ msec}$
5. Qual è il tempo medio necessario per leggere  $c = 50$  blocchi contigui?  
 $t_S + t_L + c \times t_B = 5 + 3 + 50 \times 0,020 \text{ msec} = \text{ca } 9 \text{ msec}$

**Domanda 2** Una catena di supermercati ha una base di dati dei propri clienti che dispongono di una “tessera fedeltà,” con varie informazioni su ciascun cliente, fra cui (a) il totale dei punti acquisiti attraverso l’uso della tessera e (b) il negozio della catena cui fa riferimento (ad esempio, quello presso cui ha inizialmente richiesto la tessera). Si vuole eseguire su di essa l’interrogazione che calcola, per ciascun negozio, il numero dei clienti, la somma dei punti fedeltà dei clienti e la relativa media per cliente. Indicare quale livello di isolamento (READ UNCOMMITTED, READ COMMITTED, REPEATABLE READ o SERIALIZABLE) si potrebbe scegliere in ciascuno dei seguenti casi (si supponga che, in generale, sia stato rilevato che, nel corso degli inserimenti e delle modifiche, vengono inseriti valori sbagliati anche di vari ordini di grandezza, che sono poi corretti prima del commit):

1. L’operazione è eseguita mentre vengono inseriti molti nuovi clienti, con la finalità di acquisire informazioni approssimate ma ragionevolmente indicative sugli andamenti complessivi.
2. L’operazione è eseguita mentre vengono modificati i valori dei punti fedeltà di alcuni clienti (in ciascun negozio una piccola percentuale), con la finalità di acquisire informazioni approssimate ma ragionevolmente indicative sugli andamenti complessivi.
3. L’operazione è eseguita in un momento in cui non ci sono aggiornamenti di alcun genere, con la finalità di individuare i primi tre negozi da premiare in una campagna promozionale sulla base dei punti acquisiti dai rispettivi clienti.
4. L’operazione è eseguita mentre vengono inseriti molti nuovi clienti, con la finalità di individuare i primi tre negozi da premiare in una campagna promozionale sulla base dei punti acquisiti dai rispettivi clienti.
5. L’operazione è eseguita mentre vengono modificati i valori dei punti fedeltà di alcuni clienti (in ciascun negozio una piccola percentuale), con la finalità di individuare i primi tre negozi da premiare in una campagna promozionale sulla base dei punti acquisiti dai rispettivi clienti.

Risposte				
1.	2.	3.	4.	5.
S	RC	RU	S	RR

**Domanda 3** Molti DBMS utilizzano, oltre a quelli basati sui B<sup>+</sup>-tree, anche indici basati su una tecnica detta "bitmap." Un indice bitmap, su un attributo A con v valori diversi, per una relazione con N<sub>R</sub> ennuple, è costituito da v vettori (uno per ciascun valore di A) di N<sub>R</sub> bit ciascuno (un bit per ogni ennupla della relazione): l'i-esimo bit del vettore associato al valore a<sub>j</sub> è 1 se il valore della i-esima ennupla sull'attributo A è a<sub>j</sub> e 0 se il valore è diverso. L'accesso ai vettori è organizzato per mezzo di un albero. In sostanza, rispetto ad un B<sup>+</sup>-tree, abbiamo una struttura di accesso simile, ma con foglie diverse: nel B<sup>+</sup>-tree, per ogni valore dell'attributo abbiamo una lista di indirizzi, mentre nell'indice bitmap abbiamo un vettore di bit. Esiste poi una struttura (tabella, indice o altro) che associa gli indirizzi delle ennuple ai numeri da 1 a N<sub>R</sub>. Sulla base degli elementi forniti, con riferimento ad un sistema con blocchi di dimensione D = 8 KB e indirizzi dei record di lunghezza p = 8 B e ad una relazione R con N<sub>R</sub> = 3.200.000 ennuple e (fra gli altri) quattro attributi, A, C, E, F ciascuno con v = 20 valori diversi e una lunghezza a = 6 byte

1. valutare lo spazio SBM (in blocchi) necessario per le foglie di un indice bitmap su A; valutare anche lo spazio S<sub>I</sub> necessario per una tabella (una sorta di array in memoria secondaria, ad esempio) che permetta di associare un indirizzo a ciascun valore compreso fra 1 e N<sub>R</sub>; per confronto, calcolare lo spazio ST necessario per le foglie di un B<sup>+</sup>-tree (con riferimenti ai record) per un indice su A

Poi, valutare (sia nel caso di indici bitmap, sia nel caso di B<sup>+</sup>-tree) il costo (in termini di numero di accessi a memoria secondaria, supponendo che tutti i nodi intermedi di tutti gli indici siano nei buffer; basta quindi contare il numero di foglie da visitare e l'accesso alle ennuple, quando necessario) per le seguenti interrogazioni

2. SELECT COUNT(\*) FROM R WHERE A = 5
3. SELECT \* FROM R WHERE A = 5
4. SELECT \* FROM R WHERE A = 15 AND C = 14 AND E = 13 AND F = 9 (considerare i valori indipendenti l'uno dall'altro; quindi assumere che il risultato contenga circa N<sub>R</sub>/v<sup>4</sup> ennuple)

Riportare le risposte nella tabella sottostante, nelle righe da 1 a 4, indicando formula e valore numerico.

Risposte	
	Indici B <sup>+</sup> -tree
1	$SBM = \frac{v \times N_R}{8 \times D} = \frac{20 \times 3.200.000}{8 \times 8000} = 1000$ $S_I = \frac{N_R \times p}{D} = \frac{3.200.000 \times 8}{8000} = 3200$
2	$ST = \frac{v \times a + N_R \times p}{D \times f} =$ $= \frac{20 \times 6 + 3.200.000 \times 8}{8000 \times 0,66} \approx 4800$ <p>(f è il fattore di riempimento del B<sup>+</sup>-tree assunto pari a 0,66)</p>
3	<p>Basta accedere ad un vettore:</p> $\frac{N_R}{8 \times D} = \frac{3.200.000}{8 \times 8000} = 50$
4	<p>Come sopra, più l'accesso ai record (attraverso la struttura intermedia)</p> $\frac{N_R}{8 \times D} + 2 \frac{N_R}{v} = 50 + 2 \times \frac{3.200.000}{20} \approx 320.000$
4	<p>Come sopra, più l'accesso ai record</p> $\frac{1}{a} \frac{N_R \times p}{D \times f} + \frac{N_R}{v} = \dots + \frac{3.200.000}{20} = 160.000$
4	<p>(Si suppone un indice per ciascuno degli attributi)</p> <p>Quattro volte il caso 2, più l'accesso ai record</p> $4 \times \frac{N_R}{8 \times D} + 2 \times \frac{N_R}{v^4} = 200 + 40$
4	<p>(Si suppone un indice per ciascuno degli attributi)</p> <p>Quattro volte il caso 2, più l'accesso ai record</p> $4 \times \frac{1}{a} \frac{N_R \times p}{D \times f} + \frac{N_R}{v^4} = 4 \times 240 + 40$

**Domanda 4** Specificare per ciascuno dei seguenti schedule, la sua appartenenza alle varie classi (scrivere “sì” o “no” nelle caselle della tabella; assumere che gli identificatori delle transazioni corrispondano ai timestamp).

Risposte					
	Seriale	VSR	CSR	TS	2PL
$r_2(x)w_2(x)r_1(x)w_1(x)$	seriale	VSR	CSR	no	2PL
$r_1(x)w_1(x)r_2(x)w_2(x)r_0(y)w_1(y)$	no	VSR	CSR	TS	no
$r_1(x)r_2(z)w_2(z)w_1(x)r_2(x)w_2(x)$	no	VSR	CSR	TS	2PL
$r_1(y)r_2(z)r_2(y)w_2(y)w_2(z)r_1(z)$	no	no	no	no	no

**Domanda 5** Considerare uno schema dimensionale relativo agli esami, che utilizzi, come tabella dei fatti e come una delle dimensioni, le relazioni come quelle qui schematizzate:

KStudente	KCorso	KData	Voto	...
201	301	405	26	...
201	302	406	28	...
202	301	407	30	...
202	303	407	22	...
...	...	...	...	...

KCorso	Titolo	Crediti	...
301	Fisica I	6	...
302	Chimica	9	...
303	Geometria	6	...
...	...	...	...

Supporre che si presentino le seguenti esigenze di modifica:

- i corsi cambiano nome nel tempo: per esempio, il corso nella prima ennupla potrebbe ad un certo punto cambiare nome da “Fisica I” in “Meccanica”; interessano selezioni e aggregazioni relative agli esami tanto con riferimento al nome del corso (al momento dell’esame) quanto alla sua identità (un codice che viene introdotto allo scopo, ma non sempre viene utilizzato, perché alcuni analisti preferiscono fare riferimento al nome corrente del corso); le modifiche sono rare, ma è possibile che ci siano corsi con vari cambiamenti di nome;
- per ogni corso, interessa rappresentare anche il docente, per supportare analisi sugli esami svolti da ciascun docente; i docenti cambiano nel tempo e passano da un corso all’altro (e possono anche tenere più corsi nello stesso momento, ma ogni corso ha, in un certo giorno, un solo docente); è disponibile l’informazione relativa ai docenti dei corsi nel tempo (per tutto il periodo, anche passato, di interesse).

Mostrare nuove versioni delle due tabelle che permettano di soddisfare le esigenze sopra citate (mostrare anche i dati, con riferimento a quelli presenti negli esempi sopra, aggiungendo nuovi dati ragionevoli, che permettano di comprendere le modifiche).

*Possibile soluzione*

La tabella dei fatti può essere estesa aggiungendo una dimensione supplementare, che dipende da Corso e Data e può quindi essere facilmente aggiunta anche ad una tabella dei fatti esistente (i dati, come detto, sono disponibili):

KStudente	KCorso	KData	KDocente	Voto	...
201	301	405	701	26	...
201	302	406	701	28	...
202	301	407	702	30	...
202	303	407	703	22	...
...	...	...	...	...	...

Per la dimensione può essere utile la tecnica della “slowly changing dimension,” con un nuovo elemento (quindi una ennupla nella relazione) per ogni modifica. Viste le specifiche, qui potrebbe essere utile introdurre un codice, che non cambia nel tempo, e avere due attributi per il nome, con quello attuale e quello “dell’epoca.”

KCorso	Codice	Titolo	TitoloAttuale	Crediti	...
301	FIS01	Fisica I	Meccanica	6	...
302	CHIM01	Chimica	Chimica	9	...
303	GEOM01	Geometria	Geometria	6	...
...	...	...	...	...	...
309	FIS01	Meccanica	Meccanica	6	...

## Basi di dati II, primo modulo Prova parziale — 22 aprile 2010 — Compito C Cenni sulle soluzioni

Rispondere su questo fascicolo. Consegnare anche i fogli usati come brutta copia, che però non verranno corretti.  
Tempo a disposizione: un'ora e trenta minuti.

Cognome \_\_\_\_\_ Nome \_\_\_\_\_ Matricola \_\_\_\_\_ Ordin. \_\_\_\_\_

**Domanda 1** Considerare un disco con una velocità di rotazione di  $v = 10.000$  giri al minuto e un tempo medio di posizionamento della testina (tempo di seek)  $t_S = 5$  msec. Ogni traccia contiene  $P = 240$  blocchi e ogni blocco contiene  $B = 4$  KB. Rispondere alle seguenti domande mostrando formula e valore numerico numero (N.B. non servono calcolatrici, i risultati sono semplici, approssimare  $1 \text{ MB} = 1000 \text{ KB}$ ).

1. Qual è tempo medio di latenza (attesa dovuta alla rotazione)  $t_L$ ?  
è il tempo necessario per mezzo giro:  $1/2 \times 1/v \times 60 = 1/2 \times 1/10.000 \times 60 = 0,003 \text{ sec} = 3 \text{ msec}$
2. Qual è la massima velocità di trasferimento (in MB al secondo)?  
 $v/60 \times P \times B = 10.000/60 \times 240 \times 4 \text{ KB} = 160.000 \text{ KB/sec} = 160 \text{ MB/sec}$
3. Qual è il tempo minimo  $t_B$  necessario per leggere un blocco?  
 $1/v \times 60 \times 1/P = 1/10.000 \times 60 \times 1/240 = 1/10.000 \times 1/240 \times 60 \text{ sec} = 25 \mu\text{sec}$
4. Qual è il tempo medio necessario per leggere un blocco?  
 $t_S + t_L + t_B = 5 + 3 + 0,025 \text{ msec} = \text{ca } 8 \text{ msec}$
5. Qual è il tempo medio necessario per leggere  $c = 40$  blocchi contigui?  
 $t_S + t_L + c \times t_B = 5 + 3 + 40 \times 0,025 \text{ msec} = \text{ca } 9 \text{ msec}$

**Domanda 2** Una catena di supermercati ha una base di dati dei propri clienti che dispongono di una “tessera fedeltà,” con varie informazioni su ciascun cliente, fra cui (a) il totale dei punti acquisiti attraverso l’uso della tessera e (b) il negozio della catena cui fa riferimento (ad esempio, quello presso cui ha inizialmente richiesto la tessera). Si vuole eseguire su di essa l’interrogazione che calcola, per ciascun negozio, il numero dei clienti, la somma dei punti fedeltà dei clienti e la relativa media per cliente. Indicare quale livello di isolamento (READ UNCOMMITTED, READ COMMITTED, REPEATABLE READ o SERIALIZABLE) si potrebbe scegliere in ciascuno dei seguenti casi (si supponga che, in generale, sia stato rilevato che, nel corso degli inserimenti e delle modifiche, vengono inseriti valori sbagliati anche di vari ordini di grandezza, che sono poi corretti prima del commit):

1. L’operazione è eseguita mentre vengono inseriti alcuni nuovi clienti (per ciascun negozio pochi rispetto a quelli già presenti), con la finalità di acquisire informazioni approssimate ma ragionevolmente indicative sugli andamenti complessivi.
2. L’operazione è eseguita mentre vengono modificati i valori dei punti fedeltà di tutti i clienti (a seguito di una ridefinizione dei criteri di assegnazione dei punti stessi), con la finalità di acquisire informazioni approssimate ma ragionevolmente indicative sugli andamenti complessivi.
3. L’operazione è eseguita in un momento in cui non ci sono aggiornamenti di alcun genere, con la finalità di individuare i primi tre negozi da premiare in una campagna promozionale sulla base dei punti acquisiti dai rispettivi clienti.
4. L’operazione è eseguita mentre vengono inseriti alcuni nuovi clienti (per ciascun negozio pochi rispetto a quelli già presenti), con la finalità di individuare i primi tre negozi da premiare in una campagna promozionale sulla base dei punti acquisiti dai rispettivi clienti.
5. L’operazione è eseguita mentre vengono modificati i valori dei punti fedeltà di tutti i clienti (a seguito di una ridefinizione dei criteri di assegnazione dei punti stessi), con la finalità di individuare i primi tre negozi da premiare in una campagna promozionale sulla base dei punti acquisiti dai rispettivi clienti.

Risposte				
1.	2.	3.	4.	5.
RC	RR	RU	S	RR

**Domanda 3** Molti DBMS utilizzano, oltre a quelli basati sui B<sup>+</sup>-tree, anche indici basati su una tecnica detta "bitmap." Un indice bitmap, su un attributo A con v valori diversi, per una relazione con N<sub>R</sub> ennuple, è costituito da v vettori (uno per ciascun valore di A) di N<sub>R</sub> bit ciascuno (un bit per ogni ennupla della relazione): l'i-esimo bit del vettore associato al valore a<sub>j</sub> è 1 se il valore della i-esima ennupla sull'attributo A è a<sub>j</sub> e 0 se il valore è diverso. L'accesso ai vettori è organizzato per mezzo di un albero. In sostanza, rispetto ad un B<sup>+</sup>-tree, abbiamo una struttura di accesso simile, ma con foglie diverse: nel B<sup>+</sup>-tree, per ogni valore dell'attributo abbiamo una lista di indirizzi, mentre nell'indice bitmap abbiamo un vettore di bit. Esiste poi una struttura (tabella, indice o altro) che associa gli indirizzi delle ennuple ai numeri da 1 a N<sub>R</sub>. Sulla base degli elementi forniti, con riferimento ad un sistema con blocchi di dimensione B = 8 KB e indirizzi dei record di lunghezza p = 8 B e ad una relazione R con N<sub>R</sub> = 3.200.000 ennuple e (fra gli altri) quattro attributi, A, C, E, F ciascuno con v = 20 valori diversi e una lunghezza a = 6 byte

1. valutare lo spazio SBM (in blocchi) necessario per le foglie di un indice bitmap su A; valutare anche lo spazio S<sub>I</sub> necessario per una tabella (una sorta di array in memoria secondaria, ad esempio) che permetta di associare un indirizzo a ciascun valore compreso fra 1 e N<sub>R</sub>; per confronto, calcolare lo spazio ST necessario per le foglie di un B<sup>+</sup>-tree (con riferimenti ai record) per un indice su A

Poi, valutare (sia nel caso di indici bitmap, sia nel caso di B<sup>+</sup>-tree) il costo (in termini di numero di accessi a memoria secondaria, supponendo che tutti i nodi intermedi di tutti gli indici siano nei buffer; basta quindi contare il numero di foglie da visitare e l'accesso alle ennuple, quando necessario) per le seguenti interrogazioni

2. SELECT COUNT(\*) FROM R WHERE A = 5
3. SELECT \* FROM R WHERE A = 5
4. SELECT \* FROM R WHERE A = 15 AND C = 14 AND E = 13 AND F = 9 (considerare i valori indipendenti l'uno dall'altro; quindi assumere che il risultato contenga circa N<sub>R</sub>/v<sup>4</sup> ennuple)

Riportare le risposte nella tabella sottostante, nelle righe da 1 a 4, indicando formula e valore numerico.

Risposte	
	Indici B <sup>+</sup> -tree
1	$SBM = \frac{v \times N_R}{8 \times B} = \frac{20 \times 3.200.000}{8 \times 8000} = 1000$ $S_I = \frac{N_R \times p}{B} = \frac{3.200.000 \times 8}{8000} = 3200$
2	$ST = \frac{v \times a + N_R \times p}{B \times f} =$ $= \frac{20 \times 6 + 3.200.000 \times 8}{8000 \times 0,66} \approx 4800$ <p>(f è il fattore di riempimento del B<sup>+</sup>-tree assunto pari a 0,66)</p>
3	<p>Basta accedere ad un vettore:</p> $\frac{N_R}{8 \times B} = \frac{3.200.000}{8 \times 8000} = 50$
4	<p>Come sopra, più l'accesso ai record (attraverso la struttura intermedia)</p> $\frac{N_R}{8 \times B} + 2 \frac{N_R}{v} = 50 + 2 \times \frac{3.200.000}{20} \approx 320.000$
4	<p>Come sopra, più l'accesso ai record</p> $\frac{1}{a} \frac{N_R \times p}{B \times f} + \frac{N_R}{v} = \dots + \frac{3.200.000}{20} = 160.000$
4	<p>(Si suppone un indice per ciascuno degli attributi)</p> <p>Quattro volte il caso 2, più l'accesso ai record</p> $4 \times \frac{N_R}{8 \times B} + 2 \times \frac{N_R}{v^4} = 200 + 40$
4	<p>(Si suppone un indice per ciascuno degli attributi)</p> <p>Quattro volte il caso 2, più l'accesso ai record</p> $4 \times \frac{1}{a} \frac{N_R \times p}{B \times f} + \frac{N_R}{v^4} = 4 \times 240 + 40$



**Domanda 4** Specificare per ciascuno dei seguenti schedule, la sua appartenenza alle varie classi (scrivere “sì” o “no” nelle caselle della tabella; assumere che gli identificatori delle transazioni corrispondano ai timestamp).

Risposte					
	Seriale	VSR	CSR	TS	2PL
$r_1(x)r_2(z)w_2(z)w_1(x)r_2(x)w_2(x)$	no	VSR	CSR	TS	2PL
$r_1(y)r_2(z)r_2(y)w_2(y)w_2(z)r_1(z)$	no	no	no	no	no
$r_2(x)w_2(x)r_1(x)w_1(x)$	seriale	VSR	CSR	no	2PL
$r_1(x)w_1(x)r_2(x)w_2(x)r_0(y)w_1(y)$	no	VSR	CSR	TS	no

**Domanda 5** Considerare uno schema dimensionale relativo agli esami, che utilizzi, come tabella dei fatti e come una delle dimensioni, le relazioni come quelle qui schematizzate:

KStudente	KCorso	KData	Voto	...
401	201	305	27	...
401	202	306	28	...
402	201	307	30	...
402	203	307	22	...
...	...	...	...	...

KCorso	Titolo	Crediti	...
201	Fisica I	6	...
202	Chimica	9	...
203	Geometria	6	...
...	...	...	...

Supporre che si presentino le seguenti esigenze di modifica:

- i corsi cambiano nome nel tempo: per esempio, il corso nella prima ennupla potrebbe ad un certo punto cambiare nome da “Fisica I” in “Meccanica”; interessano selezioni e aggregazioni relative agli esami tanto con riferimento al nome del corso (al momento dell’esame) quanto alla sua identità (un codice che viene introdotto allo scopo, ma non sempre viene utilizzato, perché alcuni analisti preferiscono fare riferimento al nome corrente del corso); le modifiche sono rare, ma è possibile che ci siano corsi con vari cambiamenti di nome;
- per ogni corso, interessa rappresentare anche il docente, per supportare analisi sugli esami svolti da ciascun docente; i docenti cambiano nel tempo e passano da un corso all’altro (e possono anche tenere più corsi nello stesso momento, ma ogni corso ha, in un certo giorno, un solo docente); è disponibile l’informazione relativa ai docenti dei corsi nel tempo (per tutto il periodo, anche passato, di interesse).

Mostrare nuove versioni delle due tabelle che permettano di soddisfare le esigenze sopra citate (mostrare anche i dati, con riferimento a quelli presenti negli esempi sopra, aggiungendo nuovi dati ragionevoli, che permettano di comprendere le modifiche).

*Possibile soluzione*

La tabella dei fatti può essere estesa aggiungendo una dimensione supplementare, che dipende da Corso e Data e può quindi essere facilmente aggiunta anche ad una tabella dei fatti esistente (i dati, come detto, sono disponibili):

KStudente	KCorso	KData	KDocente	Voto	...
401	201	305	701	27	...
401	202	306	701	28	...
402	201	307	702	30	...
402	203	307	703	22	...
...	...	...	...	...	...

Per la dimensione può essere utile la tecnica della “slowly changing dimension,” con un nuovo elemento (quindi una ennupla nella relazione) per ogni modifica. Viste le specifiche, qui potrebbe essere utile introdurre un codice, che non cambia nel tempo, e avere due attributi per il nome, con quello attuale e quello “dell’epoca.”

KCorso	Codice	Titolo	TitoloAttuale	Crediti	...
201	FIS01	Fisica I	Meccanica	6	...
202	CHIM01	Chimica	Chimica	9	...
203	GEOM01	Geometria	Geometria	6	...
...	...	...	...	...	...
209	FIS01	Meccanica	Meccanica	6	...

## Basi di dati II, primo modulo Prova parziale — 22 aprile 2010 — Compito D Cenni sulle soluzioni

Rispondere su questo fascicolo. Consegnare anche i fogli usati come brutta copia, che però non verranno corretti.  
Tempo a disposizione: un'ora e trenta minuti.

Cognome \_\_\_\_\_ Nome \_\_\_\_\_ Matricola \_\_\_\_\_ Ordin. \_\_\_\_\_

**Domanda 1** Considerare un disco con una velocità di rotazione di  $r = 10.000$  giri al minuto e un tempo medio di posizionamento della testina (tempo di seek)  $t_S = 5$  msec. Ogni traccia contiene  $N = 300$  blocchi e ogni blocco contiene  $B = 4$  KB. Rispondere alle seguenti domande mostrando formula e valore numerico numero (N.B. non servono calcolatrici, i risultati sono semplici, approssimare  $1 \text{ MB} = 1000 \text{ KB}$ ).

1. Qual è tempo medio di latenza (attesa dovuta alla rotazione)  $t_L$ ?  
è il tempo necessario per mezzo giro:  $1/2 \times 1/r \times 60 = 1/2 \times 1/10.000 \times 60 = 0,003 \text{ sec} = 3 \text{ msec}$
2. Qual è la massima velocità di trasferimento (in MB al secondo)?  
 $r/60 \times N \times B = 10.000/60 \times 300 \times 4 \text{ KB} = 200.000 \text{ KB/sec} = 200 \text{ MB/sec}$
3. Qual è il tempo minimo  $t_B$  necessario per leggere un blocco?  
 $1/r \times 60 \times 1/N = 1/10.000 \times 60 \times 1/300 = 1/10.000 \times 1/300 \times 60 \text{ sec} = 20 \mu\text{sec}$
4. Qual è il tempo medio necessario per leggere un blocco?  
 $t_S + t_L + t_B = 5 + 3 + 0,020 \text{ msec} = \text{ca } 8 \text{ msec}$
5. Qual è il tempo medio necessario per leggere  $n = 50$  blocchi contigui?  
 $t_S + t_L + n \times t_B = 5 + 3 + 50 \times 0,020 \text{ msec} = \text{ca } 9 \text{ msec}$

**Domanda 2** Una catena di supermercati ha una base di dati dei propri clienti che dispongono di una “tessera fedeltà,” con varie informazioni su ciascun cliente, fra cui (a) il totale dei punti acquisiti attraverso l’uso della tessera e (b) il negozio della catena cui fa riferimento (ad esempio, quello presso cui ha inizialmente richiesto la tessera). Si vuole eseguire su di essa l’interrogazione che calcola, per ciascun negozio, il numero dei clienti, la somma dei punti fedeltà dei clienti e la relativa media per cliente. Indicare quale livello di isolamento (READ UNCOMMITTED, READ COMMITTED, REPEATABLE READ o SERIALIZABLE) si potrebbe scegliere in ciascuno dei seguenti casi (si supponga che, in generale, sia stato rilevato che, nel corso degli inserimenti e delle modifiche, vengono inseriti valori sbagliati anche di vari ordini di grandezza, che sono poi corretti prima del commit):

1. L’operazione è eseguita mentre vengono modificati i valori dei punti fedeltà di alcuni clienti (in ciascun negozio una piccola percentuale), con la finalità di acquisire informazioni approssimate ma ragionevolmente indicative sugli andamenti complessivi.
2. L’operazione è eseguita mentre vengono modificati i valori dei punti fedeltà di tutti i clienti (a seguito di una ridefinizione dei criteri di assegnazione dei punti stessi), con la finalità di acquisire informazioni approssimate ma ragionevolmente indicative sugli andamenti complessivi.
3. L’operazione è eseguita mentre vengono inseriti molti nuovi clienti, con la finalità di individuare i primi tre negozi da premiare in una campagna promozionale sulla base dei punti acquisiti dai rispettivi clienti.
4. L’operazione è eseguita mentre vengono modificati i valori dei punti fedeltà di alcuni clienti (in ciascun negozio una piccola percentuale), con la finalità di individuare i primi tre negozi da premiare in una campagna promozionale sulla base dei punti acquisiti dai rispettivi clienti.
5. L’operazione è eseguita in un momento in cui non ci sono aggiornamenti di alcun genere, con la finalità di acquisire informazioni approssimate ma ragionevolmente indicative sugli andamenti complessivi.

Risposte				
1.	2.	3.	4.	5.
RC	RR	S	RR	RU

**Domanda 3** Molti DBMS utilizzano, oltre a quelli basati sui B<sup>+</sup>-tree, anche indici basati su una tecnica detta “bitmap.” Un indice bitmap, su un attributo A con v valori diversi, per una relazione con N<sub>R</sub> ennuple, è costituito da v vettori (uno per ciascun valore di A) di N<sub>R</sub> bit ciascuno (un bit per ogni ennupla della relazione): l’i-esimo bit del vettore associato al valore a<sub>j</sub> è 1 se il valore della i-esima ennupla sull’attributo A è a<sub>j</sub> e 0 se il valore è diverso. L’accesso ai vettori è organizzato per mezzo di un albero. In sostanza, rispetto ad un B<sup>+</sup>-tree, abbiamo una struttura di accesso simile, ma con foglie diverse: nel B<sup>+</sup>-tree, per ogni valore dell’attributo abbiamo una lista di indirizzi, mentre nell’indice bitmap abbiamo un vettore di bit. Esiste poi una struttura (tabella, indice o altro) che associa gli indirizzi delle ennuple ai numeri da 1 a N<sub>R</sub>. Sulla base degli elementi forniti, con riferimento ad un sistema con blocchi di dimensione D = 4 KB e indirizzi dei record di lunghezza p = 8 B e ad una relazione R con N<sub>R</sub> = 1.600.000 ennuple e (fra gli altri) quattro attributi, A, C, E, F ciascuno con v = 20 valori diversi e una lunghezza a = 6 byte

1. valutare lo spazio SBM (in blocchi) necessario per le foglie di un indice bitmap su A; valutare anche lo spazio S<sub>I</sub> necessario per una tabella (una sorta di array in memoria secondaria, ad esempio) che permetta di associare un indirizzo a ciascun valore compreso fra 1 e N<sub>R</sub>; per confronto, calcolare lo spazio ST necessario per le foglie di un B<sup>+</sup>-tree (con riferimenti ai record) per un indice su A

Poi, valutare (sia nel caso di indici bitmap, sia nel caso di B<sup>+</sup>-tree) il costo (in termini di numero di accessi a memoria secondaria, supponendo che tutti i nodi intermedi di tutti gli indici siano nei buffer; basta quindi contare il numero di foglie da visitare e l’accesso alle ennuple, quando necessario) per le seguenti interrogazioni

2. SELECT COUNT(\*) FROM R WHERE A = 5
3. SELECT \* FROM R WHERE A = 5
4. SELECT \* FROM R WHERE A = 15 AND C = 14 AND E = 13 AND F = 9 (considerare i valori indipendenti l’uno dall’altro; quindi assumere che il risultato contenga circa N<sub>R</sub>/v<sup>4</sup> ennuple)

Riportare le risposte nella tabella sottostante, nelle righe da 1 a 4, indicando formula e valore numerico.

Risposte	
	Indici B <sup>+</sup> -tree
1	$SBM = \frac{v \times N_R}{8 \times D} = \frac{20 \times 1.600.000}{8 \times 4000} = 1000$ $S_I = \frac{N_R \times p}{D} = \frac{1.600.000 \times 8}{4000} = 3200$
2	$ST = \frac{v \times a + N_R \times p}{D \times f} =$ $= \frac{20 \times 6 + 1.600.000 \times 8}{4000 \times 0,66} \approx 4800$ <p>(f è il fattore di riempimento del B<sup>+</sup>-tree assunto pari a 0,66)</p>
3	<p>Basta accedere ad un vettore:</p> $\frac{N_R}{8 \times D} = \frac{1.600.000}{8 \times 4000} = 50$
4	<p>Come sopra, più l’accesso ai record (attraverso la struttura intermedia)</p> $\frac{N_R}{8 \times D} + 2 \times \frac{N_R}{v} = 50 + 2 \times \frac{1.600.000}{20} \approx 160.000$
4	<p>Come sopra, più l’accesso ai record</p> $\frac{1}{a} \frac{N_R \times p}{D \times f} + \frac{N_R}{v} = \dots + \frac{1.600.000}{20} = 80.000$
4	<p>(Si suppone un indice per ciascuno degli attributi)</p> <p>Quattro volte il caso 2, più l’accesso ai record</p> $4 \times \frac{N_R}{8 \times D} + 2 \times \frac{N_R}{v^4} = 200 + 20$
4	<p>(Si suppone un indice per ciascuno degli attributi)</p> <p>Quattro volte il caso 2, più l’accesso ai record</p> $4 \times \frac{1}{a} \frac{N_R \times p}{D \times f} + \frac{N_R}{v^4} = 4 \times 240 + 20$

**Domanda 4** Specificare per ciascuno dei seguenti schedule, la sua appartenenza alle varie classi (scrivere “sì” o “no” nelle caselle della tabella; assumere che gli identificatori delle transazioni corrispondano ai timestamp).

Risposte					
	Seriale	VSR	CSR	TS	2PL
$r_2(x)w_2(x)r_1(x)w_1(x)$	seriale	VSR	CSR	no	2PL
$r_1(y)r_2(z)r_2(y)w_2(y)w_2(z)r_1(z)$	no	no	no	no	no
$r_1(x)w_1(x)r_2(x)w_2(x)r_0(y)w_1(y)$	no	VSR	CSR	TS	no
$r_1(x)r_2(z)w_2(z)w_1(x)r_2(x)w_2(x)$	no	VSR	CSR	TS	2PL

**Domanda 5** Considerare uno schema dimensionale relativo agli esami, che utilizzi, come tabella dei fatti e come una delle dimensioni, le relazioni come quelle qui schematizzate:

KStudente	KCorso	KData	Voto	...	KCorso	Titolo	Crediti	...
401	301	105	28	...	301	Fisica I	6	...
401	302	106	28	...	302	Chimica	9	...
402	301	107	30	...	303	Geometria	6	...
402	303	107	22	...	...	...	...	...
...	...	...	...	...				

Supporre che si presentino le seguenti esigenze di modifica:

- i corsi cambiano nome nel tempo: per esempio, il corso nella prima ennupla potrebbe ad un certo punto cambiare nome da “Fisica I” in “Meccanica”; interessano selezioni e aggregazioni relative agli esami tanto con riferimento al nome del corso (al momento dell’esame) quanto alla sua identità (un codice che viene introdotto allo scopo, ma non sempre viene utilizzato, perché alcuni analisti preferiscono fare riferimento al nome corrente del corso); le modifiche sono rare, ma è possibile che ci siano corsi con vari cambiamenti di nome;
- per ogni corso, interessa rappresentare anche il docente, per supportare analisi sugli esami svolti da ciascun docente; i docenti cambiano nel tempo e passano da un corso all’altro (e possono anche tenere più corsi nello stesso momento, ma ogni corso ha, in un certo giorno, un solo docente); è disponibile l’informazione relativa ai docenti dei corsi nel tempo (per tutto il periodo, anche passato, di interesse).

Mostrare nuove versioni delle due tabelle che permettano di soddisfare le esigenze sopra citate (mostrare anche i dati, con riferimento a quelli presenti negli esempi sopra, aggiungendo nuovi dati ragionevoli, che permettano di comprendere le modifiche).

*Possibile soluzione*

La tabella dei fatti può essere estesa aggiungendo una dimensione supplementare, che dipende da Corso e Data e può quindi essere facilmente aggiunta anche ad una tabella dei fatti esistente (i dati, come detto, sono disponibili):

KStudente	KCorso	KData	KDocente	Voto	...
401	301	105	701	28	...
401	302	106	701	28	...
402	301	107	702	30	...
402	303	107	703	22	...
...	...	...	...	...	...

Per la dimensione può essere utile la tecnica della “slowly changing dimension,” con un nuovo elemento (quindi una ennupla nella relazione) per ogni modifica. Viste le specifiche, qui potrebbe essere utile introdurre un codice, che non cambia nel tempo, e avere due attributi per il nome, con quello attuale e quello “dell’epoca.”

KCorso	Codice	Titolo	TitoloAttuale	Crediti	...
301	FIS01	Fisica I	Meccanica	6	...
302	CHIM01	Chimica	Chimica	9	...
303	GEOM01	Geometria	Geometria	6	...
...	...	...	...	...	...
309	FIS01	Meccanica	Meccanica	6	...