

Informatica

Andrea Pollini

Il tema proposto quest'anno si compone di una prima parte in cui viene chiesto di analizzare l'evoluzione del mercato discografico con l'introduzione della digitalizzazione, soffermandosi su opportunità e problemi che si sono aperti in seguito a questa rivoluzione. Gli aspetti di questa rivoluzione nel processo di gestione dei brani musicali sono molteplici e sotto gli occhi di tutti. Dapprima l'introduzione a vasta scala di formati di digitalizzazione ottimizzata e compressa, come il formato mp3 hanno consentito di poter conservare un numero elevato di canzoni su di un supporto rispetto ad un CD audio tradizionale; infatti un file mp3 è di almeno un ordine di grandezza più piccolo della classica traccia audio CD. Inoltre esistono strumenti per convertire un file mp3 in un file audio CD, per cui il processo, a meno della piccola fetta di informazioni audio perse nel processo di compressione e decompressione, risulta invertibile. Questo formato compresso, da quando è stato supportato in lettori portatili, ha consentito una diffusione a macchia d'olio della musica. Inoltre l'avvento di programmi di condivisione della musica ha, da un lato, aperto sia la possibilità per gruppi non famosi di farsi conoscere, sia il poter acquistare un brano musicale per meno di un euro, direttamente con un click del mouse. Purtroppo, dall'altra faccia della medaglia la facilità di condivisione della musica ha aperto la strada anche allo scambio illecito di materiale protetto da copyright.

Per risolvere il tema d'esame proposto occorre procedere innanzitutto analizzando tutto il testo, ed andando ad estrarre le informazioni che ci consentiranno come primo passo di realizzare il modello concettuale del nostro sistema informativo. Dallo schema concettuale realizzeremo lo schema logico, aggiungendo eventuali attributi che possono rivelarsi utili per la risoluzione dei vari quesiti. Successivamente definiremo il codice SQL che implementa il sistema informativo richiesto ed andremo a scrivere le interrogazioni SQL che rispondono ai quesiti della traccia. Per realizzare un corretto modello concettuale occorre ricordare che risulta fondamentale non tanto descrivere la realtà del sistema di pubblicazione di una casa editrice in genere, quanto piuttosto di realizzare il nostro sistema informativo seguendo una modalità cosiddetta *goal oriented*. Nel modello concettuale andremo a definire le entità del sistema, ovve-

ro gli oggetti reali o astratti che ci servono per rispondere ai quesiti proposti, senza aggiungere complessità e strutture non necessarie.

LA TRACCIA MINISTERIALE

INFORMATICA GENERALE ED APPLICAZIONI GESTIONALI Indirizzo: PROGRAMMATORI

Le tecniche di digitalizzazione, trasmissione, memorizzazione e riproduzione di brani musicali hanno rivoluzionato il mercato discografico creando nuove opportunità e nuovi problemi. Il candidato inizi la sua prova soffermandosi brevemente su quelli che considera i principali aspetti di tale fenomeno e, successivamente, analizzi e risolva il seguente problema:

Una casa editrice pubblica riviste prevalentemente rivolte a specialisti di varie discipline operanti in ambiente scientifico. Sapendo che:

- le pubblicazioni, che possono essere quindicinali, mensili, bimestrali, semestrali, sono diverse tra loro per l'ambito disciplinare a cui sono rivolte e per il prezzo unitario di ciascuna;
- sono inviate ad abbonati che abbiano sottoscritto un abbonamento di durata annuale o triennale;
- sono distribuite, su ordinazione, ad alcuni punti vendita dislocati all'interno di istituti di ricerca e di alta formazione italiani e stranieri;

A) definire e documentare adeguatamente la struttura di una base di dati che possa rispondere ad interrogazioni del tipo:

1. elenco con nome, cognome e telefono degli abbonati di genere femminile residenti in Spagna;
2. elenco con la data di scadenza degli abbonamenti stipulati in giugno 2009 e con il nominativo e l'indirizzo di posta elettronica dei relativi abbonati;
3. elenco con la denominazione e il recapito degli istituti francesi che, nel 2008, hanno emesso ordini per la fornitura di riviste bimestrali di Chimica industriale;
4. numero complessivo delle riviste ordinate dal Politecnico di Torino nell'anno in corso.

B) descrivere le caratteristiche del software che si intende utilizzare per la realizzazione della base di dati;

C) spiegare nei dettagli come effettuare almeno una delle quattro interrogazioni proposte dalla traccia;

D) realizzare una pagina web, scegliendola a piacere tra quelle che potrebbero essere ragionevolmente presenti in un ipotetico sito della casa editrice.

I dati mancanti possono essere opportunamente scelti dal candidato.

ESAMI CONCLUSIVI

Modello concettuale

Invece di una descrizione generale della realtà, andremo a descrivere un modello che della realtà in esame possieda solo quegli attributi e caratteristiche necessari a risolvere i nostri quesiti. Descriveremo le entità in gioco e di ogni entità delincheremo quali siano i suoi attributi essenziali, successivamente descriveremo le eventuali relazioni tra le entità del modello concettuale, necessarie per risolvere i quesiti del tema che stiamo affrontando.

Dovremo creare un modello di una catena di distribuzione di una casa editrice, che vende riviste secondo diverse modalità che andremo via via ad analizzare.

L'entità di base sarà la *Rivista*. Di ogni rivista sappiamo che avrà un titolo ed un ambito disciplinare di riferimento (ad es. matematica, fisica, chimica). Inoltre ogni rivista avrà una periodicità, ovvero ogni quanti giorni viene pubblicata. Nella traccia viene indicato che ogni rivista avrà anche un proprio prezzo. Chiaramente è vero nella realtà, ma nel nostro caso il prezzo non serve per risolvere nessuno dei quesiti proposti, per cui possiamo evitare di includerlo nel modello concettuale.

Le riviste sono distribuite secondo due tipi di canali di distribuzione. Il primo è la spedizione in abbonamento annuale o triennale ed è fatta a singoli abbonati. Per prima cosa vediamo come organizzare il modello concettuale dell'entità *Abbonato* che contiene i dati anagrafici di un singolo abbonato. Questi dati saranno esterni all'entità *Abbonamento*, che invece conterrà i dati di un singolo abbonamento in quanto potrebbe capitare che un abbonato richieda più abbonamenti a riviste diverse, secondo una relazione N-N.

Gli attributi che nel nostro modello descrivono un *Abbonato* sono nome, cognome, sesso, telefono, email e stato di residenza. La relazione che descrive un *Abbonamento* sarà invece individuata dalla durata in anni (1 o 3) e dal mese ed anno di stipula. È importante separare anno e mese in quanto viene richiesto in uno dei quesiti di calcolare la data di scadenza dell'abbonamento stesso e mantenendo separati i due dati il calcolo risulta più facile.

L'altro canale di distribuzione è quello su ordinazione. Per descrivere concettualmente questo canale di distribuzione utilizzeremo un'entità, *Istituto*. A questa entità affiancheremo la relazione *Ordina* che descriverà l'ordine effettuato a una rivista, che sarà l'altro capo della relazione.

L'entità *Istituto* fa riferimento agli istituti di ricerca e di alta formazione, italiani e stranieri, che possono ordinare le riviste che la casa editrice pubblica. Gli attributi dell'entità

Istituto saranno la denominazione dell'istituto, il suo indirizzo e lo stato di appartenenza.

La relazione *Ordina* descriverà invece l'ordine che un istituto può fare di alcune riviste, e sarà caratterizzata dall'anno di ordinazione, dal riferimento alla rivista ordinata e dal numero di copie della stessa. Per semplicità possiamo assumere che se un istituto ordina più riviste diverse, queste siano in ordini diversi. La relazione ha cardinalità N-N in quanto ogni istituto potrà ordinare più riviste ed ogni rivista potrà essere ordinata da più istituti.

Progettazione della base dati

Realizzato il modello concettuale ora possiamo concentrarci sulla creazione della base dati. La traduzione del modello concettuale in quello logico e nella conseguente creazione del codice SQL per le tabelle che faranno parte della base dati è un'operazione abbastanza metodica. Entità e relazioni definite nel modello concettuale definiscono le tabelle della base dati.

Nella creazione del modello logico partiamo da quanto ottenuto dal modello concettuale, ma introduciamo ulteriori informazioni quali ad esempio il tipo di dato (numerico, stringa) relativo ad ogni attributo ed eventuali attributi aggiuntivi che serviranno nella base dati per individuare univocamente una riga (chiavi primarie) nelle tabelle o effettuare il riferimento ad una riga di una tabella esterna (chiavi esterne) nelle relazioni ove serve.

L'entità *Rivista* nel modello logico verrà utilizzata al plurale e verrà aggiunto un attributo numerico autoincrementante che servirà nella base dati ad individuare la singola rivista. Nel modello logico avremo per questo una tabella *riviste* con i seguenti campi:

Tabella	<i>riviste</i>	
Campo	Tipo dato	vincoli
id	SERIAL	NOT NULL PRIMARY KEY
titolo	VARCHAR(120)	NOT NULL
ambito	VARCHAR(120)	NOT NULL
periodicità	VARCHAR(120)	NOT NULL

Le colonne, tranne *id* che è numerico, sono stringhe non nulle, in quanto campi vuoti non avrebbero senso per questa tabella. Il campo *id* è intero autoincrementante, indicato con SERIAL. Inoltre *id* è una chiave primaria e il fatto di essere autoincrementante assicura che tutte le righe avranno sulla colonna *id* valori distinti.

L'entità *Istituto* diventerà nel modello logico la tabella *istituti* con le colonne indicate nel seguente schema:

Tabella	istituti	
Campo	Tipo dato	vincoli
id	SERIAL	NOT NULL PRIMARY KEY
denominazione	VARCHAR(120)	NOT NULL
ambito	VARCHAR(120)	NOT NULL
periodicità	VARCHAR(120)	NOT NULL

La terza ed ultima entità dello schema concettuale è l'entità *Abbonato*. Il suo corrispettivo nello schema logico sarà la tabella *abbonati* così definita:

Tabella	abbonati	
Campo	Tipo dato	vincoli
id	SERIAL	NOT NULL PRIMARY KEY
nome	VARCHAR(120)	NOT NULL
cognome	VARCHAR(120)	NOT NULL
sexo	CHAR(1)	NOT NULL
telefono	VARCHAR(120)	NOT NULL
email	VARCHAR(120)	NOT NULL
stato	VARCHAR(120)	NOT NULL

Ora occorre creare il modello logico delle relazioni. Le relazioni del modello concettuale che abbiamo costruito dovranno avere oltre agli attributi già evidenziati, anche dei riferimenti, le cosiddette chiavi esterne, alle chiavi primarie delle entità cui fanno riferimento.

La relazione *Abbonamento* nel modello logico corrisponderà alla tabella *abbonamenti*. Oltre ai campi del modello concettuale dovremo inserire due campi interi che realizzino il riferimento a riviste ed abbonati. Questi campi vengono solitamente chiamati secondo la convenzione di utilizzare il nome della tabella di riferimenti in minuscolo, seguito dai caratteri '_id', per indicare che è un riferimento ad una tabella esterna. Essendo una tabella di relazione, non è necessario possieda una chiave primaria. Quindi lo schema logico sarà:

Tabella	abbonamenti	
Campo	Tipo dato	vincoli
abbonato_id	INTEGER	NOT NULL
rivista_id	INTEGER	NOT NULL
mese	VARCCHAR(3)	NOT NULL
anno	INTEGER	NOT NULL
durata	INTEGER	NOT NULL

ESAMI CONCLUSIVI

L'ultima relazione da inserire nel modello logico è *Ordina*. Tale relazione corrisponde alla tabella *ordini*, seguendo lo schema utilizzato per la relazione abbonamento.

Tabella	ordini	
Campo	Tipo dato	vincoli
istituto_id	INTEGER	NOT NULL
rivista_id	INTEGER	NOT NULL
copie	INTEGER	NOT NULL
anno	INTEGER	NOT NULL

Creazione della base dati

Creato il modello logico, sarà ora facile scrivere le istruzioni SQL che consentiranno di realizzare la base dati fino ad ora definita. Chiameremo *casa_editrice* il database e la definizione completa dello stesso in linguaggio SQL sarà la seguente:

```
CREATE DATABASE casa_editrice;

CREATE TABLE riviste
(
    titolo character varying(120) NOT NULL,
    ambito character varying(120) NOT NULL,
    periodicità character varying(120) NOT NULL,
    id serial NOT NULL PRIMARY KEY
);

CREATE TABLE abbonamenti (
    abbonato_id integer NOT NULL,
    rivista_id integer NOT NULL,
    mese character varying(3) NOT NULL,
    anno integer NOT NULL,
    durata integer NOT NULL
);

CREATE TABLE abbonati (
    nome character varying(120) NOT NULL,
    cognome character varying(120) NOT NULL,
    sesso character(1) NOT NULL,
    telefono character varying(120) NOT NULL,
    stato character varying(120) NOT NULL,
    email character varying(120) NOT NULL,
    id SERIAL NOT NULL
);

CREATE TABLE istituti (
    id SERIAL NOT NULL,
    denominazione character varying(120) NOT NULL,
    indirizzo character varying(120) NOT NULL,
    stato character varying(120) NOT NULL
);
```

ESAMI CONCLUSIVI

```
CREATE TABLE ordini (
  istituto_id integer NOT NULL,
  rivista_id integer NOT NULL,
  anno integer NOT NULL,
  copie integer NOT NULL
);

CREATE TABLE riviste (
  titolo character varying(120) NOT NULL,
  ambito character varying NOT(120) NULL,
  periodicit  character varying(120) NOT NULL,
  id SERIAL NOT NULL
);
```

Le interrogazioni SQL

La prima interrogazione che viene richiesta è quella relativa all'elenco con nome, cognome e telefono degli abbonati di genere femminile residenti in Spagna. Questa interrogazione è facile, infatti, è sufficiente impostare una query SQL che estragga i dati delle colonne necessarie dalla tabella abbonati, aggiungendo i filtri relativi a sesso e stato di residenza. Poiché le due condizioni dovranno essere verificate entrambe, dobbiamo unire i due filtri con una clausola AND. Quindi la query richiesta sarà

```
SELECT
  nome,cognome,telefono
FROM
  abbonati
WHERE
  sesso='F' AND stato='SPAGNA'
```

La seconda query è l'elenco con la data di scadenza degli abbonamenti stipulati in giugno 2009 e con il nominativo e l'indirizzo di posta elettronica dei relativi abbonati. Questa query è più complessa della precedente. La realizzeremo per passaggi successivi. Visto l'assunto di memorizzare nella tabella abbonamenti mese ed anno di stipula e durata dello stesso, partiamo estraendo solo le date di scadenza degli abbonamenti. Dovremo quindi aggiungere all'anno di stipula gli anni di durata, che saranno uno o tre.

```
SELECT
  mese,
  anno+durata AS anno_scadenza
FROM abbonamenti
```

Come secondo passaggio estendiamo la query affinché mostri oltre ai dati già estratti anche il nome, cognome e l'indirizzo email degli abbonati. Quindi aggiungeremo la condizione *abbonati.id=abbonamenti.abbonato_id* che realizza la *join* tra le due tabelle, estrapolando dalla tabella abbonati i dati necessa-

ri. A questa, in AND, aggiungeremo le condizioni relative all'anno ed al mese di stipula. La query completa sarà:

```
SELECT
  mese,
  anno+durata AS anno_scadenza,
  nome,cognome,email
FROM abbonamenti,abbonati
WHERE abbonati.id=abbonamenti.abbonato_id
AND
  mese = 'GIU'
AND
  anno=2009
```

La terza richiesta è l'elenco con la denominazione e il recapito degli istituti francesi che, nel 2008, hanno emesso ordini per la fornitura di riviste bimestrali di Chimica industriale. Per prima cosa estrapoliamo l'elenco degli istituti francesi, selezionando le colonne di interesse

```
SELECT
  denominazione,
  indirizzo
FROM
  istituti
WHERE
  stato='FRANCIA'
```

ora effettuiamo le *join* con le tabelle ordini e riviste, aggiungendo le clausole di filtro. Utilizzeremo la clausola DISTINCT in quanto la query potrebbe restituire righe duplicate. La query risultante diventa

```
SELECT DISTINCT
  denominazione,
  indirizzo
FROM
  istituti,ordini,riviste
WHERE
  stato='FRANCIA'
AND
  istituti.id=ordini.istituto_id
AND
  riviste.id=ordini.rivista_id
AND riviste.periodicit ='BIMESTRALE'
AND riviste.ambito='CHIMICA INDUSTRIALE'
AND ordini.anno='2008'
```

4) "numero complessivo delle riviste ordinate dal Politecnico di Torino nell'anno in corso".



```

SELECT
  SUM(ordini.copie)
FROM
  ordini,istituti
WHERE
  istituti.id=ordini.istituto_id
AND
  istituti.denominazione='Politecnico di Torino'
AND
  anno='2008'

```

Struttura dell'applicazione

La struttura dell'applicazione che implementi il sistema informativo della casa editrice dovrebbe seguire tutte le tecniche di buona progettazione che sono state definite dall'informatica degli ultimi anni. In particolar modo utilizzare un paradigma di tipo MVC (Model-View-Controller) consente di realizzare applicazioni più flessibili ed aperte a future evoluzioni del prodotto. Realizzando una completa separazione tra la gestione dei dati e la presentazione degli stessi, infatti, si aumenta l'espandibilità dell'applicazione.

Il Model è la componente che si occupa di presentare i dati memorizzati nel database in un modo specifico per il dominio applicativo. Ad esempio nel nostro caso il *Model* presenterà una funzione per ogni query richiesta, nascondendo alle altre parti dell'applicazione come la query viene effettuata. Modificare la query nel *Model* non modifica nessun'altra parte dell'applicazione, visto che viene modificato solo il funzionamento interno della funzione del *model*, e non i parametri o il tipo dati ritornato.

La *View* è il componente che visualizza all'utente i dati secondo quanto richiesto, consentendo l'interazione con l'utente. Una particolarità è che per un singolo *Model* si possono avere più *View* diverse per scopi diversi.

Il *Controller* gestisce l'interazione con l'utente e si occupa di aggiornare i *Model* e di conseguenza le *View*. Esiste infatti un meccanismo tipico di MVC secondo cui una *View* viene "avvisata" dei cambiamenti al *Model* di riferimento. Tuttavia questa parte non è essenziale ad esempio in un'applicazione web dove un aggiornamento della pagina comporta un *refresh* automatico dei dati.

Nel nostro caso il *Model* potrebbe essere chiamato *ModelCasaEditrice*. Esso esporrà quattro funzioni che chiameremo *query1*, *query2*, *query3* e *query4*. Questi nomi non sono esplicativi in effetti, tuttavia in questo caso vista la complessità delle query saremmo stati costretti ad utilizzare nomi lunghissimi. In generale sarebbe meglio utilizzare nomi esplicativi. Ad esempio *numero_riviste_politecnico_torino_2008*, oppure, creare delle funzioni dando possibilità di usare i parametri nella costruzione delle query.

Nel contesto MVC che abbiamo introdotto, realizzare una pagina web non è altro se non progettare una vista della nostra applicazione. Per dare consistenza alle pagine del sito defini-

ESAMI CONCLUSIVI

remo un *layout* comune a tutte con una struttura del tipo indicato dalla seguente tabella.

header	
menù	contenuto
footer	

Nell'*header* metteremo titolo e *link* alle sezioni principali, comuni a tutte le pagine del sito. Nel *footer* staranno i *link* ad eventuali pagine

con informazioni di *privacy* e di *copyright*, anch'esse comuni a tutto il sito. La parte del menù consentirà di selezionare quali funzionalità richiedere e noi potremo concentrarci solo sulla parte del contenuto. Le quattro query che abbiamo prima costruito potranno essere rappresentate in modo uniforme in una tabella. Lo pseudocodice necessario per realizzarle, assumendo che le funzioni del *model* ritornino un *array* associativo con chiave il nome della colonna e con valore il valore relativo alla riga in oggetto, fornirà la seguente funzione:

```

def stampa_tabella_query(array)
  colonne = [x for x in array.keys() if 'id' not in x]
  print "<table>"
  print "<th>"
  for col in colonne:
    print "<td>"
    print col
    print "</td>"
  print "</th>"
  for row in array:
    print "<tr>"
    for col in colonne:
      print "<td>"
      print row[col]
      print "</td>"
    print "</tr>"
  print "</table>"

```

In questo modo stamperemo a video la tabella relativa alla query effettuata. Potremo poi completare la pagina con altre informazioni che riterremo interessanti.

La traccia proposta quest'anno è molto varia, consentendo allo studente di effettuare più scelte di quali aspetti approfondire del problema proposto. Delle quattro interrogazioni proposte ne viene richiesta almeno una, probabilmente questa scelta deriva dal fatto che la parte discorsiva iniziale potrebbe occupare un po' troppo tempo se lo studente non sa gestirsi al meglio. La progettazione della base dati non è semplicissima, ma procedendo con ordine, prima realizzando lo schema ER e poi traducendolo nel modello logico e nella conseguente implementazione SQL, si procede abbastanza spediti. Le query vanno affrontate per passi, come abbiamo visto, e solo la terza rappresenta difficoltà di rilievo.

Andrea Pollini

Dip. di Matematica e Fisica «N. Tartaglia»
Università Cattolica del Sacro Cuore, Brescia