

TUTORATO DI BASI DI DATI

Mattia Fazzi: mattia.fazzi@student.unife.it

02/04/2019

MODELLO RELAZIONALE

- Il modello dei dati relazionale è basato sul concetto di relazione.
- Una relazione è una tabella di valori:
 - Che può essere pensata come un insieme di righe o colonne.
 - Ciascuna riga rappresenta un fatto corrispondente all'istanza di un'entità o di un'associazione del minimondo.
 - Ciascuna riga possiede un elemento o un insieme di elementi che la identificano univocamente (chiave).

ESEMPIO

- Relazione Film:
- Attributi: Titolo, Data Pubblicazione, Descrizione, Codice.
- Codice è la chiave del film e potrebbe a scelta di lunghezza fissa: in questo caso abbiamo un dominio di valori ben definito.
- Il dominio legato all'attributo di un entità rappresenta tutti i possibili valori validi per un attributo. Es: codice potrebbe essere un valore numerico di sei cifre. Dominio da 000000 – 999999.
- Data di pubblicazione ha un formato di rappresentazione ben definito dal dd-mm-yyyy.

TUPLA

- La tupla è un insieme ordinato di valori che rappresentano un'entità.
- I valori devono rispettare il dominio dei relativi attributi.
- ES Film:
- <<«Il signore degli anelli – La compagnia dell'anello, 18-02-2002, «Nella terra di mezzo vivono...»,123456»
- L'insieme delle tuple della stessa entità rappresentano la nostra relazione e le colonne della relazione sono gli attributi.

DEFINIZIONI

- Per chiarezza si possono utilizzare termini informali per indicare i concetti di un database.
- Relazione -> Tabella
- Attributo/Dominio -> Colonna
- Tupla -> Riga
- Dominio -> Valori di una colonna
- Schema di relazione -> Definizione di una tabella
- Estensione (Stato) -> Tabella popolata

CARATTERISTICHE DELLE RELAZIONI

- L'ordinamento di una tabella non è obbligatorio: di solito la tabella viene mostrata nell'ordine in cui vengono inserite nuove tuple all'interno di essa oppure ordinata secondo la chiave.
- Ci possono essere valori NULL per alcuni attributi: NULL è un valore particolare e specifica che quel particolare attributo non è ancora stato definito.

VINCOLI DI INTEGRITA'

- I vincoli sono condizioni che devono essere rispettati.
- Esistono tre tipi di vincoli:
 - Vincoli sulla chiave;
 - Vincoli di integrità dell'entità;
 - Vincoli di integrità referenziale.

VINCOLI SULLA CHIAVE

- Per vincolo di chiave si intende l'impossibilità di avere due chiavi con lo stesso valore.
- Es: Se nella tabella film decidiamo che codice è chiave dell'entità, non ci potranno essere due film con lo stesso codice.
- La chiave deve essere univoca per poter identificare univocamente la tupla nell'insieme delle tuple della stessa relazione.

VINCOLI DI INTEGRITA' SULLE ENTITA'

- Per vincolo di integrità sulle entità si intende l'impossibilità che una chiave possa avere valore NULL.
- Come detto precedentemente, la chiave identifica univocamente l'entità e non può essere NULL.
- Per scelte progettistiche, si può decidere di avere questo vincolo anche su altri attributi della stessa entità anche se non sono chiave.
- Es: Possiamo decidere di non permettere a un film di avere valore NULL dato che un film senza un titolo avrebbe poco senso.

VINCOLO DI INTEGRITA' REFERENZIALE

- Il vincolo di integrità referenziale coinvolge due relazioni: la relazione referenziante e la relazione riferita.
- Tuple nella relazione referenziante R1 hanno attributi FK (Foreign Key) che puntano a PK (Private Key) di una relazione R2.
- E' importante che tutti i valori della colonna/attributo che riferiscono alle chiavi primarie della relazione R2 abbiano una corrispondenza.
- Detta in parole povere: non è possibile che un valore della FK di una relazione R1 abbia un valore che punta a R2 che non esiste.

VINCOLO DI INTEGRITA' SEMANTICO

- Basati sulla semantica dell'applicazione: vincoli difficilmente esprimibili nel modello ma importanti a livello applicativo e progettuale.
- Es: il numero massimo di ore, per dipendente, per tutti i progetti su cui lavora deve essere 56 ore.
- In questo caso non si ha un vincolo su un attributo singolo ma il vincolo è sulla somma delle ore: il risultato della somma delle ore non è salvato da nessuna parte ma sarà il risultato di una funzione.

VINCOLO DI INTEGRITA' SEMANTICO

- SQL consente l'uso di trigger e asserzioni per controllare che tali vincoli vengano rispettati: triggers e asserzioni vengono «invocati» in seguito a particolari azioni ad esempio inserimento, cancellazione o modifica di una tupla.

OPERAZIONI: SELECT

- SELECT: l'operazione di select consente di selezionare un insieme di tuple di una relazione che soddisfano una determinata condizione.

$$\sigma_{\langle \text{condizione di selezione} \rangle}(\mathbf{R})$$

- SIGMA specifica il tipo di operazione (in questo caso di select)
- Tra < > va inserita la condizione di selezione
- R è il nome della relazione su cui stiamo facendo la selezione

PROPRIETA' DELLA SELECT

- Schema della relazione risultate uguale allo schema della relazione originaria.
- Proprietà commutativa
- Operazioni in cascata possono essere eseguite in qualunque ordine.
- Operazioni in cascata possono essere sostituite da una select unica con condizione la congiunzione logica delle condizioni delle select.

OPERAZIONI: PROIEZIONE

- Ci consente di selezionare solamente alcuni attributi di una data relazione. L'operazione è vista come una partizione verticale della relazione.

$$\pi_{\langle \text{elenco attributi} \rangle}(\mathbf{R})$$

- Il pigreco specifica l'operazione di proiezione
- Tra < > si inseriscono gli attributi desiderati
- R è il nome della relazione su cui facciamo la proiezione

OPERAZIONI: RIDENOMINAZIONE

- Consente di nominare il risultato di un operazione intermedia in modo da poter utilizzarle la relazione in un operazione successiva.

$$\text{DEP5_EMPS} \leftarrow \sigma_{\text{DNO}=5}(\text{EMPLOYEE})$$

- In questo caso si sta creando una relazione DEP5_EMPS che, a partire da EMPLOYEE, seleziona solamente gli impiegati con DNO = 5.

OPERAZIONI INSIEMISTICHE

- Supponendo di avere due relazioni $R1 (A1, A2, \dots)$ e $R2(B1, B2, \dots)$ in cui $\text{dom}(A_i) = \text{dom}(B_i)$ (compatibili all'unione), possiamo definire le seguenti operazioni:
 - Unione: la relazione risultante avrà tutte le tuple della relazione $R1$ e $R2$ non duplicate;
 - Intersezione: la relazione risultante avrà tutte le tuple in comune tra le due relazioni;
 - Sottrazione: la relazione definita dall'operazione $R1 - R2$ avrà come risultato tutte le tuple di $R1$ che non sono presenti in $R2$.

PRODOTTO CARTESIANO

- Date due relazioni $R1$ e $R2$ non compatibili all'unione, l'operazione «prodotto cartesiano» tra $R1$ ed $R2$ ($R1 \times R2$) restituisce una relazione $R3$ che per tuple le combinazioni tra le tuple delle due relazioni.

JOIN

- Un'operazione di prodotto cartesiano seguita da una selezione sulle tuple che sono correlate delle due relazione corrisponde all'operazione di join.
- La join è una delle operazioni più importanti all'interno del database perché permette di mettere in relazione due relazioni tramite un'associazione.

JOIN

- L'operazione di join viene fatta sugli attributi di due relazioni.

$$\mathbf{R} \bowtie_{\langle \text{condizione di join} \rangle} \mathbf{S}$$

- All'interno della condizione di join possiamo trovare più condizioni unite tra loro tramite congiunzione (AND)

AGENZIA DI VIAGGI

Un'agenzia di viaggi vi commissiona il database per la loro piattaforma online su cui i clienti (nome, cognome, sesso, codfiscale, data nascita, data iscrizione) possono contattare direttamente un agente di viaggio (nome, cognome, sesso, data nascita, data assunzione, codfiscale) tramite messaggi (testo, data invio) per chiedere informazioni riguardo a offerte del sito. Le offerte (costo, data creazione, durata di validità, durata della vacanza) sono pacchetti composti da: un pernottamento in un hotel (nome, città, indirizzo, #stelle, email hotel) o in B&B(nome, città, indirizzo, email proprietario, nome proprietario) e modalità di viaggio se previsto (mezzo di trasporto, città partenza, città arrivo, durata del viaggio, costo). L'agente di viaggio deve completare la prenotazione(data prenotazione, data partenza, data ritorno, costo complessivo) dell'offerta da parte di un cliente.

OPERAZIONI

- Selezionare le offerte che costano almeno 200 euro.
- Selezionare il nome e indirizzo dei soggiorni che si trovano a firenze.
- Selezionare la data partenza e data ritorno delle prenotazioni che hanno costo complessivo minore di 350 euro e che sono state prenotate dopo il 2016.
- Selezionare il nome degli agenti di viaggio che hanno effettuato una prenotazione con costo complessivo maggiore di 400 euro.
- Selezionare le offerte che comprendono soggiorni che si trovano a milano.

PALESTRA

Il CUS vi incarica di progettare il nuovo database della centro sportivo. Ogni utente (nome, cognome, data nascita, sesso, codice fiscale) sarà un cliente (data iscrizione) o un personal trainer (data assunzione). Ogni cliente può iscriversi a determinati corsi (nome, tipologia, descrizione) che saranno tenuti da un personal trainer. Ogni cliente può prenotare una seduta di allenamento che sarà supervisionata da un personal trainer. La prenotazione avrà data, orario e durata della seduta. Ogni utente può essere iscritto a un determinato club sportivo (tipologia e nome allenatore) che si terrà in una specifica palestra (nome, città, indirizzo). Ogni utente può decidere di comprare un abbonamento (identificativo, durata, data inizio, data fine) o di comprare un numero generico di entrate (numero entrate: 10,20 o 30, data acquisto, data validità).

OPERAZIONI

Selezionare gli utenti che si sono iscritti dopo settembre 2017.

Selezionare il nome e il cognome dei personal trainer femmine.

Selezionare le sedute di allenamento con durata superiore a 60 minuti.

Selezionare il nome e il cognome dei personal trainer che hanno tenuto una seduta di allenamento prima delle ore 15.

Selezionare l'indirizzo delle palestre che sono sede del club allenato da «Ciccio Graziani».

Selezionare il nome, cognome e sesso dei personal trainer tengono un corso a cui è iscritto almeno un cliente che ha un abbonamento che finisce dopo gennaio 2019.

RAGGRUPPAMENTO

Contare quanti maschi hanno sottoscritto un abbonamento da 10 ingressi.

Trovare il nome e cognome del personal trainer più giovane.

Calcolare la media dell'età delle persone iscritta al corso di tipologia «Crossfit»