

Basi di Dati – Modulo 1

Prof. De Marsico

18 gennaio 2022

1) Sia data la seguente base di dati relativa ad una compagnia aerea

AEREO(<u>ID</u> , Costruttore, Modello, NPosti, AnnoC, AnnoR)
VOLO(<u>Sigla</u> , Partenza, Arrivo, Orario)
AEROPORTO(<u>Sigla</u> , Città)
VIAGGIO(<u>ID</u> , SiglaVolo, Aereo, Data)
PERSONALE(<u>CE</u> , ID, Nome, Cognome, DataNascita)
EQUIPAGGIO(<u>Viaggio</u> , <u>Pers</u>)

NOTE

- AEREO.AnnoC e AEREO.AnnoR sono rispettivamente l'anno di costruzione e **l'anno di ultima revisione di un aereo** (se l'aereo non è stato mai revisionato AnnoR=0000)
- VOLO.Sigla è la sigla di un volo
- VOLO.Partenza e VOLO.Arrivo sono **sigle** di aeroporti
- Le città più grandi possono avere più aeroporti
- Un volo viene effettuato sempre **alla stessa ora in più date**
- VIAGGIO.SiglaVolo è la sigla di un volo effettuato in data VIAGGIO.Data
- VIAGGIO.Aereo è **l'ID di un aereo**
- EQUIPAGGIO.Viaggio è **l'ID di un viaggio**
- EQUIPAGGIO.Pers è **l'ID di un membro del personale**
- Un membro del personale può partecipare a più viaggi in date diverse.

1a) Trovare i modelli di aerei costruiti prima del 2010 e mai revisionati con cui sono stati effettuati voli diretti a New York tra il 2019 e il 2021

1b) Trovare nome e cognome e Codice Fiscale di membri del personale che non hanno mai fatto parte dell'equipaggio di voli partiti da Roma dopo le ore 23:00.

2) Siano dati lo schema $R=ABCDEF$ e l'insieme di dipendenze funzionali

$F=\{A \rightarrow CD, BC \rightarrow AE, D \rightarrow BE, AB \rightarrow EF\}$

2a) Determinare la/le chiavi dello schema

2b) Dire se lo schema è 3NF e giustificare l'affermazione

2c) Trovare una decomposizione dello schema tale che:

- ogni sottoschema è 3NF
 - la decomposizione preserva le dipendenze
 - la decomposizione ha un join senza perdita.
-

3) E' dato un file di 135.700 record. Ogni record occupa 370 byte. Un blocco contiene 2048 byte. Un puntatore a blocco occupa 5 byte. Si utilizza una organizzazione hash con record distribuiti uniformemente tra 200 bucket.

3a) Calcolare l'occupazione in blocchi della bucket directory e dei bucket

3b) Calcolare il costo medio di una ricerca considerando chiavi di hash univoche

3c) Quanti bucket occorrerebbero per avere un costo medio di ricerca minore o uguale a 20 accessi?

1) Sia data la seguente base di dati relativa ad una compagnia aerea

AEREO(ID, Costruttore, Modello, NPosti, AnnoC, AnnoR)
 VOLO(Sigla, Partenza, Arrivo, Orario)
 AEROPORTO(Sigla, Città)
 VIAGGIO(ID, SiglaVolo, Aereo, Data)
 PERSONALE(CF, ID, Nome, Cognome, DataNascita)
 EQUIPAGGIO(Viaggio, Pers)

NOTE

- AEREO.AnnoC e AEREO.AnnoR sono rispettivamente l'anno di costruzione e l'anno di ultima revisione di un aereo (se l'aereo non è stato mai revisionato AnnoR=0000)
- VOLO.Sigla è la sigla di un volo
- VOLO.Partenza e VOLO.Arrivo sono **sigle** di aeroporti
- Le città più grandi possono avere più aeroporti
- Un volo viene effettuato sempre **alla stessa ora in più date**
- VIAGGIO.SiglaVolo è la sigla di un volo effettuato in data VIAGGIO.Data
- VIAGGIO.Aereo è l'**ID di un aereo**
- EQUIPAGGIO.Viaggio è l'**ID di un viaggio**
- EQUIPAGGIO.Pers è l'**ID di un membro del personale**
- Un membro del personale può partecipare a più viaggi in date diverse.

1a) Trovare i modelli di aerei costruiti prima del 2010 e mai revisionati con cui sono stati effettuati voli diretti a New York tra il 2019 e il 2021

1b) Trovare nome e cognome e Codice Fiscale di membri del personale che non hanno mai fatto parte dell'equipaggio di voli partiti da Roma dopo le ore 23:00.

1a)

$$\text{Aereo nat R} = \sigma_{\text{Aereo}} \left(\begin{array}{l} \text{AnnoC} < 2010 \\ \text{AnnoR} = \infty \text{ or } 0000 \end{array} \right)$$

$$\text{Viaggio NY} = \pi_{\text{Aereo}} \left(\sigma_{\text{Aereo}} \left(\begin{array}{l} \text{SiglaVolo} = \text{SiglaVolo} \\ \text{Aereo} = \text{Sigla} \end{array} \right) \left(\begin{array}{l} \text{Città} = \text{"New York"} \\ \text{Orario} < 23:00 \text{ or } 23:00 \end{array} \right) \right)$$

$$\text{out} = \pi_{\text{Modello}} \left(\text{Viaggio NY} \text{ nat R} \right)$$

1b)

$$\text{Viaggio DA R} = \pi_{\text{Viaggio ID}} \left(\sigma_{\text{Viaggio ID}} \left(\begin{array}{l} \text{Data} > 23:00 \\ \text{Città} = \text{Roma} \end{array} \right) \left(\begin{array}{l} \text{SiglaVolo} = \text{SiglaVolo} \\ \text{Data} > 23:00 \end{array} \right) \right)$$

$$\text{Pers Viaggio DA R} = \left(\text{Viaggio DA R} \text{ nat Equipaggio nat Personale} \right)$$

$$\text{out} = \pi_{\text{CF, Nome, Cognome}} \left(\text{Personale - Personale Viaggio DA R} \right)$$

2) Siano dati lo schema $R=ABCDEF$ e l'insieme di dipendenze funzionali

$F=\{A \rightarrow CD, BC \rightarrow AE, D \rightarrow BE, AB \rightarrow EF\}$

2a) Determinare la/le chiavi dello schema

2b) Dire se lo schema è 3NF e giustificare l'affermazione

2c) Trovare una decomposizione dello schema tale che:

- ogni sottoschema è 3NF
- la decomposizione preserva le dipendenze
- la decomposizione ha un join senza perdita.

2A) $(A)^+_F = ACDBEF \vee \rightarrow A \text{ è CHIAVE}$

$(BC)^+_F = BCDAEF \vee \rightarrow BC \text{ è CHIAVE}$

$(DC)^+_F = DCBEAF \rightarrow DC \text{ è CHIAVE}$

2B) Non in 3NF * $D \rightarrow BE$ poiché D non è PK e BE non PK.

2C) $I^0 F = \{ A \rightarrow C, A \rightarrow D, BC \rightarrow A, BC \rightarrow E, D \rightarrow B, D \rightarrow E, AB \rightarrow E, AB \rightarrow F \}$

I^0 Computa Minore

- $BC \rightarrow A$

$(B)^+_F = B \nrightarrow A$

$(C)^+_F = C \nrightarrow A$

- $BC \rightarrow E$

$(B)^+_F = B \nrightarrow E$

$(C)^+_F = C \nrightarrow E$

- $AB \rightarrow E$

$(A)^+_F = ACDBFE \supset E \Rightarrow \cancel{AB \rightarrow E}$

- $AB \rightarrow F$

$(A)^+_F = ACDBE \nrightarrow F$

$(B)^+_F = D \nrightarrow F$

Alb fine del 2° passo

$$F = \{ A \rightarrow C, A \rightarrow D, BC \rightarrow A, BC \rightarrow E, D \rightarrow B, D \rightarrow E, A \rightarrow E, AB \rightarrow F \}$$

III° Cerco ridondanze

- $A \rightarrow C$

$$(A)^+_{F \rightarrow A \rightarrow C} = ADCBE \not\subset C \Rightarrow \text{NON TOLGO}$$

- $A \rightarrow D$

$$(A)^+_{F \rightarrow A \rightarrow D} = ACE \not\subset D \Rightarrow \text{NON TOLGO}$$

- $BC \rightarrow A$

$$(BC)^+_{F \rightarrow BC \rightarrow A} = BCE \not\subset A \Rightarrow \text{NON TOLGO}$$

- $BC \rightarrow E$

$$(BC)^+_{F \rightarrow BC \rightarrow E} = BCAE \supset E \Rightarrow \text{TOLGO}$$

$$F = \{ A \rightarrow C, A \rightarrow D, BC \rightarrow A, \cancel{BC \rightarrow E}, D \rightarrow B, D \rightarrow E, A \rightarrow E, AB \rightarrow F \}$$

- $D \rightarrow B$

$$(D)^+_{F \rightarrow D \rightarrow B} = DE \not\subset B \Rightarrow \text{NON TOLGO}$$

- $D \rightarrow E$

$$(D)^+_{F \rightarrow D \rightarrow E} = DB \not\subset E \Rightarrow \text{NON TOLGO}$$

- $A \rightarrow E$

$$(A)^+_{F \rightarrow A \rightarrow E} = ACDBE \supset E \Rightarrow \text{TOLGO}$$

$$F = \{ A \rightarrow C, A \rightarrow D, BC \rightarrow A, \cancel{BC \rightarrow E}, D \rightarrow B, D \rightarrow E, \cancel{A \rightarrow E}, AB \rightarrow F \}$$

- $AB \rightarrow F$

$$(AB)^+_{F \rightarrow AB \rightarrow F} = ABCDE \not\subset F \Rightarrow \text{NON TOLGO}$$

Dopo 3° passo

$$F = \{ A \rightarrow C, A \rightarrow D, BC \rightarrow A, D \rightarrow B, D \rightarrow E, AB \rightarrow F \}$$

↓ UNO DETERMINANTI 2 DETERMINATI

$$\rho = \{ AC, AD, BCA, DB, DE, ABF \}$$

POICHE' SONO PRESENTI IN ρ TUTTI GLI ELEMENTI DI R

ρ PRESERVA F e σ IN $3NF$, POICHE' PRESENTI CHIACCI
(ES $A \subset AC$) HA JOIN SENZA PERDITA

3) E' dato un file di 135.700 record. Ogni record occupa 370 byte. Un blocco contiene 2048 byte. Un puntatore a blocco occupa 5 byte. Si utilizza una organizzazione hash con record distribuiti uniformemente tra 200 bucket.

3a) Calcolare l'occupazione in blocchi della bucket directory e dei bucket

3b) Calcolare il costo medio di una ricerca considerando chiavi di hash univoche

3c) Quanti bucket occorrerebbero per avere un costo medio di ricerca minore o uguale a 20 accessi?

3A)

$$N^{\circ} \text{Blocchi Bucket Dir} = \frac{\lceil N^{\circ} \text{Bucket} \rceil}{\text{Record} \times \text{Block dir}} = \frac{\lceil \frac{2048}{5} \rceil}{1} = 2$$

$$\text{Record} \times \text{Block dir} = \left\lfloor \frac{\text{Block size}}{P_{\text{size}}} \right\rfloor = \left\lfloor \frac{2048}{5} \right\rfloor = 409$$

$$\text{Block} \times \text{Bucket} = \frac{\lceil \text{Record} \times \text{Bucket} \rceil}{\text{Record} \times \text{Block}} = \frac{679}{5} = 136$$

$$\text{Record} \times \text{Bucket} = \frac{\lceil N^{\circ} \text{Record} \rceil}{N^{\circ} \text{Bucket}} = \frac{\lceil 135700 \rceil}{200} = 679$$

$$\text{Record} \times \text{Block} = \left\lfloor \frac{\text{Block size} - \text{Pointer size}}{\text{Record size}} \right\rfloor = \left\lfloor \frac{2048 - 5}{370} \right\rfloor = 5$$

3b) $\text{Avg Time} = \frac{\lceil \frac{136}{2} \rceil}{2} = 68 \text{ Accessi}$

3c) $\text{Avg Time} \leq 20$

$$\text{Block} \times \text{Bucket} \leq 40$$

$$\frac{\text{Record} \times \text{Bucket}}{\text{Record} \times \text{Block}} \leq 40 \rightarrow$$

$$\text{Record} \times \text{Bucket} \leq 40 \cdot 6$$

$$\text{Record} \times \text{Bucket} \leq 240$$

$$N^{\circ} \text{Bucket} \geq \frac{N^{\circ} \text{Record}}{240} \geq 678$$

$$\frac{N^{\circ} \text{Record}}{N^{\circ} \text{Bucket}} \leq 240$$

$$N^{\circ} \text{Record} \leq 240 \cdot N^{\circ} \text{Bucket}$$