

Basi di Dati – Modulo 1

23 gennaio 2023 - Traccia B

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 che hanno avuto l'ultima revisione più di 10 anni dopo l'anno di costruzione e con cui sono stati effettuati voli diretti a Parigi nel 2021 e 2022

1b) Trovare nome e cognome e Codice Fiscale di membri del personale che non hanno mai volato con arrivo in città diverse da Madrid.

2a) Dati lo schema di relazione $R=ABCDE$, l'insieme di dipendenze funzionali

$F=\{AB \rightarrow C, AB \rightarrow D, B \rightarrow E, CD \rightarrow E\}$ e la decomposizione $\rho=\{ABC, CDE\}$ di R ,

dire se ρ preserva F e illustrare il procedimento seguito per giungere alla risposta

2b) Dati lo schema di relazione $R=ABCDEG$, l'insieme di dipendenze funzionali

$F=\{G \rightarrow AB, A \rightarrow E, E \rightarrow B, BE \rightarrow G\}$ e la decomposizione $\rho=\{ACD, ABE, CDG\}$ di R ,

dire se ρ ha un join senza perdita e illustrare il procedimento seguito per giungere alla risposta

3) E' dato un file di 234.700 record. Ogni record occupa 357 byte. Un blocco contiene 2048 byte. Un puntatore a blocco occupa 4 byte. Si utilizza una organizzazione hash con record distribuiti uniformemente tra 350 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 15 accessi?

Esercizio 1

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(CE, 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 che hanno avuto l'ultima revisione più di 10 anni dopo l'anno di costruzione e con cui sono stati effettuati voli diretti a Parigi nel 2021 e 2022

1b) Trovare nome e cognome e Codice Fiscale di membri del personale che non hanno mai volato con arrivo in città diverse da Madrid.

1a) Anno_Rev - 10: σ (Aereo)

AnnoR > AnnoC + 10

Voli_Parigi = $\pi_{AEREO} \left(\begin{array}{c} \sigma \\ \text{01/01/2021} \\ \text{11} \\ \text{DATA} \\ \text{11} \\ \text{31/12/2022} \\ \text{11} \\ \text{Arrivo = "Parigi"} \end{array} \right) \left(\begin{array}{c} \text{Viaggio} \\ \text{Volo} \end{array} \right)$

sigla = sigla volo

out = $\pi_{modello} \left(\begin{array}{c} \text{Voli_Parigi} \\ \text{Anno_Rev_10} \end{array} \right)$

Aereo = id

1b) Aereo_polic_M = $\pi_{sigla} \left(\begin{array}{c} \sigma \\ \text{città} \\ \text{"Madrid"} \end{array} \right) \left(\begin{array}{c} \text{AEROPORTO} \end{array} \right)$

N_volo_M = $\pi_{sigla} \left(\begin{array}{c} \text{Volo} \\ \text{Aereo_polic_M} \end{array} \right)$

sigla = sigla

$$V_i \text{ New M} = \pi_{10} \left(V_{\text{Viaggio}} \bowtie V_{\text{NoVolc M}} \right)$$

SiglaVolc = Sigla

$$Equip \text{ No M} = \left(V_i \text{ New M} \bowtie_{V_{\text{Viaggio}} = 10} Equip \text{ gto} \right)$$

$$Pers \text{ No M} = \pi_{\substack{\text{Nome,} \\ \text{Cognome,} \\ \text{CF}}} \left(P_{\text{Personale}} \bowtie_{ID = pers} Equip \text{ No M} \right)$$

$$out = P_{\text{Personale}} - Pers \text{ No M}$$

Correzione Prof.

$$V_{\text{Volc}} - \text{No M MADRID} = \sigma_{\substack{\text{AEROPORTO} \neq \text{MADRID}}} (V_{\text{Volc}} \bowtie V_{\text{Viaggio}})$$

$$V_{\text{Viaggio}}^{\text{MAL}} = (V_{\text{Viaggio}} \bowtie V_{\text{Volc}} - \text{MADRID})$$

$$2b) \quad P_{\text{ERS}} - \text{MADRID} = \pi_{\text{ERS}} \left(V_{\text{Viaggio}}^{\text{MAL}} \bowtie_{\substack{\text{Viaggio-MADRID. id} \\ = \\ \text{Viaggio}}} E_{\text{QUIPAGGIO}} \right)$$

$$P_{\text{ERS}} - \text{VOLAT} = \pi_{\text{ERS}} (E_{\text{QUIPAGGIO}})$$

$$\text{MAL MADRID} = P_{\text{ERS}} - \text{VOLAT} - P_{\text{ERS}} - \text{MADRID}$$

$$\text{Sol} = \pi_{\substack{\text{Nome,} \\ \text{Cognome,} \\ \text{CF}}} \left(P_{\text{PERSONALE}} \bowtie_{ID = PERS} \text{MAL MADRID} \right)$$

Esercizio 2

2a) Dati lo schema di relazione $R=ABCDE$, l'insieme di dipendenze funzionali

$F=\{AB \twoheadrightarrow C, AB \twoheadrightarrow D, B \twoheadrightarrow E, CD \twoheadrightarrow E\}$ e la decomposizione $\rho=\{ABC, CDE\}$ di R ,

dire se ρ preserva F e illustrare il procedimento seguito per giungere alla risposta

2b) Dati lo schema di relazione $R=ABCDEG$, l'insieme di dipendenze funzionali

$F=\{G \twoheadrightarrow AB, A \twoheadrightarrow E, E \twoheadrightarrow B, BE \twoheadrightarrow G\}$ e la decomposizione $\rho=\{ACD, ABE, CDG\}$ di R ,

dire se ρ ha un join senza perdita e illustrare il procedimento seguito per giungere alla risposta

2a) dare controllo:

- $AB \twoheadrightarrow D$

- $B \twoheadrightarrow E$

tratto l'algoritmo

Parto da $AB \twoheadrightarrow D$

$Z_0 = AB$

$$S_0 = (AB \cap ABC)_F^+ \cap ABC \cup (AB \cap CDE)_F^+ \cap CDE$$

$$= (AB)_F^+ \cap ABC \cup \emptyset$$

$$= ABCDE \cap ABC \cup \emptyset = ABC$$

$S_0 \neq Z_0 \Rightarrow$ continuo

$Z_1 = ABC$

$$S_1 = (ABC \cap ABC)_F^+ \cap ABC \cup (ABC \cap CDE)_F^+ \cap CDE =$$

$$= (ABC)_F^+ \cap ABC \cup (C)_F^+ \cap CDE =$$

$$= ABCDE \cap ABC \cup C \cap CDE =$$

$$= ABC \cup C = ABC$$

$S_1 \subset Z_1 \Rightarrow m_1 \text{ feno}$

ABC now CONTINUE $D \Rightarrow AB \rightarrow D$ now

PDE SPURATA

2B

	A	B	C	D	E	G
ACD	a	b ₁	a	a	b ₁	b ₁
ABE	a	a	b ₂	b ₁	a	b ₂
CDG	b ₃	b ₃	a	a	b ₃	a

1° caso

$G \rightarrow AB$ ✓

$A \rightarrow E$ ✓

$E \rightarrow B$ ✓

$BE \rightarrow G$ ✓

	A	B	C	D	E	G
ACD	a	b₁ [ⓐ]	a	a	b₁ [ⓐ]	b₁ [ⓐ]
ABE	a	a	b ₂	b ₁	a	b₂ [ⓐ]
CDG	b ₃	b ₃	a	a	b ₃	a

2° caso

$G \rightarrow AB$ ✓

$A \rightarrow E$ ✓

$E \rightarrow B$ ✓

$BE \rightarrow G$ ✓

m_1 FERMO perché non ho CAMBIO NIENTE

⇓

non HA JOU SEUAI PERDITA

Esercizio ③

3) E' dato un file di 234.700 record. Ogni record occupa 357 byte. Un blocco contiene 2048 byte. Un puntatore a blocco occupa 4 byte. Si utilizza una organizzazione hash con record distribuiti uniformemente tra 350 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 15 accessi?

3a)

$$N^{\circ} \text{Blocchi} \times \text{Bucket Dir} = \frac{\lceil N^{\circ} \text{Bucket} \rceil}{\text{Record} \times \text{Block Size}} = \frac{\lceil \frac{234700}{350} \rceil}{512} = 1$$

$$\text{Blocchi} \times \text{Bucket} = \frac{\lceil \text{Record} \times \text{Block Size} \rceil}{\text{Record} \times \text{Block Size}} = \frac{\lceil \frac{234700 \times 357}{2048} \rceil}{512} = 135$$

$$\text{Record} \times \text{Block Dir} = \frac{\text{Block Size}}{\text{P size}} = \frac{2048}{4} = 512$$

$$\text{Record} \times \text{Block} = \frac{\text{Block Size} - \text{P size}}{\text{Record Size}} = \frac{2048 - 4}{357} = \frac{2044}{357} = 5$$

$$\text{Record} \times \text{Bucket} = \frac{\lceil N^{\circ} \text{Record} \rceil}{N^{\circ} \text{Bucket}} = \frac{\lceil \frac{234700}{350} \rceil}{350} = 671$$

3b)

$$\text{Avg Time} = \frac{\lceil \text{Blocchi} + \text{Bucket} \rceil}{2} = \frac{\lceil \frac{135}{2} \rceil}{2} = 68$$

3c)

$$\frac{\text{Blocchi} + \text{Bucket}}{2} \leq 15 \Rightarrow \text{Blocchi} \times \text{Bucket} \leq 30$$

$$\frac{\text{Record} \times \text{Bucket}}{\text{Record} \times \text{Block}} \leq 30$$

$$\text{Record} \times \text{Bucket} \leq 30 \cdot \text{Record} \times \text{Block}$$

$$\text{Read} + \text{Bucket} \leq 150$$

$$\frac{\text{N}^\circ \text{Read}}{\text{N}^\circ \text{Bucket}} \leq 150$$

$$\text{N}^\circ \text{Read} \leq 150 \cdot \text{N}^\circ \text{Bucket}$$

$$\text{N}^\circ \text{Bucket} \geq \frac{\text{N}^\circ \text{Read}}{150}$$

$$\text{N}^\circ \text{Bucket} \geq \frac{(239700)}{150}$$

$$\text{N}^\circ \text{Bucket} \geq 1598$$