

## Basi di Dati – Modulo 1

16 Ottobre 2023

1) Dato il seguente schema di una **base di dati contenente dati relativi a magazzini**

PRODOTTI(Codice,Descrizione,Prezzo\_unit)

MAGAZZINI(Codice,Indirizzo, Città)

SCORTE(Codice-P,Codice-M,N-pezzi)

NOTE:

**SCORTE** contiene i dati relativi alla presenza (**N-pezzi  $\geq 1$ , quindi NON CI SONO tuple di SCORTE con N-pezzi=0**) dei prodotti nei vari magazzini

esprimere in **algebra relazionale** le seguenti interrogazioni:

1a) Per ogni prodotto con costo unitario maggiore di 50 euro e del quale sono presenti più di 150 pezzi in almeno un magazzino di Milano, si desidera conoscere: i dati del prodotto, codice e indirizzo dei magazzini in cui sono presenti più di 150 pezzi del prodotto, e il numero dei pezzi disponibili in ognuno di questi magazzini.

1b) Restituire i dati dei magazzini in cui per tutti i prodotti presenti le scorte sono di più di 2000 pezzi o il prezzo unitario è maggiore di 500

---

2) Dato il seguente schema  $R = ABCDEH$  sul quale è definito il seguente insieme di dipendenze funzionali:

$F = \{A \rightarrow C, AB \rightarrow D, ABH \rightarrow CE, BH \rightarrow AD, D \rightarrow C, DE \rightarrow H, E \rightarrow B\}$

2a) trovare le 4 chiavi dello schema motivando la risposta\*

2b) dire se lo schema è in 3NF motivando la risposta\*

2c) trovare una decomposizione dello schema\*, in modo tale che ogni sottoschema sia in 3NF, e che la decomposizione preservi  $F$  e abbia un join senza perdita.

(\* scrivere le definizioni formali degli elementi teorici a cui si fa riferimento e lo pseudo codice degli algoritmi utilizzati)

---

3) E' dato un file di 168.600 record. Ogni record occupa 426 byte, di cui 32 per la chiave. Un blocco contiene 4096 byte. Un puntatore a blocco occupa 5 byte. Si utilizza una organizzazione ISAM

3a) Calcolare l'occupazione in blocchi del file principale considerando che tutti i blocchi contengono il massimo numero di record che consente di avere almeno il 20% di spazio libero

3b) Calcolare l'occupazione in blocchi dell'indice considerando i blocchi indice completamente pieni

3c) Calcolare il costo massimo di una ricerca

Esercizio 1

1) Dato il seguente schema di una **base di dati contenente dati relativi a magazzini**

PRODOTTI(Codice, Descrizione, Prezzo\_unit)

MAGAZZINI(Codice, Indirizzo, Città)

SCORTE(Codice-P, Codice-M, N-pezzi)

NOTE:

**SCORTE** contiene i dati relativi alla presenza (**N-pezzi**  $\geq 1$ , quindi **NON CI SONO** tuple di **SCORTE** con **N-pezzi=0**) dei prodotti nei vari magazzini

esprimere in **algebra relazionale** le seguenti interrogazioni:

1a) Per ogni prodotto con costo unitario maggiore di 50 euro e del quale sono presenti più di 150 pezzi in almeno un magazzino di Milano, si desidera conoscere: i dati del prodotto, codice e indirizzo dei magazzini in cui sono presenti più di 150 pezzi del prodotto, e il numero dei pezzi disponibili in ognuno di questi magazzini.

1b) Restituire i dati dei magazzini in cui per tutti i prodotti presenti le scorte sono di più di 2000 pezzi o il prezzo unitario è maggiore di 500

1a)

$$\begin{aligned} \text{Prodotti}_2 &= \pi_{\text{Descrizione, N-pezzi, codice}} \left( \sigma_{\text{Prezzo\_unit} > 50} \left( \text{PRODOTTI} \bowtie_{\text{codice} = \text{codice P}} \text{SCORTE} \right) \right) \\ \text{Prod in M}_1 &= \pi_{\text{codice, indirizzo}} \left( \sigma_{\text{città} = \text{"Milano"}} \left( \text{MAGAZZINI} \bowtie_{\text{codice} = \text{codice M}} \text{PRODOTTI}_2 \right) \right) \end{aligned}$$

1b)

$$\begin{aligned} \text{SCORTE}_{<2000} &= \sigma_{\substack{\text{N-pezzi} < \\ 2000}} (\text{SCORTE}) \\ \text{Prezzo\_Basso} &= \sigma_{\substack{\text{Prezzo\_unit} < \\ 500}} (\text{PRODOTTI}) \end{aligned}$$

$$S_2 = \text{SCORTE} - \text{SCORTE}_{<2000}$$

$$S_3 = \text{SCORTE} - (\text{SCORTE} \bowtie_{\text{codice P} = \text{codice M}} \text{Prezzo\_Basso})$$

$$\text{OUT} = S_2 \cup S_3$$

$$\text{FINALE} = \left( \text{MAGAZZINI} \bowtie_{\text{codice} = \text{codice M}} \text{OUT} \right)$$

## Esercizio 2

2) Dato il seguente schema  $R = ABCDEH$  sul quale è definito il seguente insieme di dipendenze funzionali:

$F = \{A \rightarrow C, AB \rightarrow D, ABH \rightarrow CE, BH \rightarrow AD, D \rightarrow C, DE \rightarrow H, E \rightarrow B\}$

2a) trovare le 4 chiavi dello schema motivando la risposta\*

2b) dire se lo schema è in 3NF motivando la risposta\*

2c) trovare una decomposizione dello schema\*, in modo tale che ogni sottoschema sia in 3NF, e che la decomposizione preservi  $F$  e abbia un join senza perdita.

(\* scrivere le definizioni formali degli elementi teorici a cui si fa riferimento e lo pseudo codice degli algoritmi utilizzati)

2a)

$$1 (AE)_F^+ = AE \subset BDH \quad E' \text{ CHIAVE}$$

$$2 (DE)_F^+ = DE \subset CBA \quad E' \text{ CHIAVE}$$

$$3 (EH)_F^+ = EH \subset ADC \quad E' \text{ CHIAVE}$$

$$4 (BH)_F^+ = BH \subset ADC \quad E' \text{ CHIAVE}$$

2b) Non è in 3NF per  $A \rightarrow C$

Perché  $A$  non è primario e  $C$  non è primo

2c)  $F = \{A \rightarrow C, AB \rightarrow D, ABH \rightarrow CE, BH \rightarrow AD, D \rightarrow C, DE \rightarrow H, E \rightarrow B\}$

$$F = \{A \rightarrow C, AB \rightarrow D, ABH \rightarrow C, ABH \rightarrow E, BH \rightarrow A, BH \rightarrow D, D \rightarrow C, DE \rightarrow H, E \rightarrow B\}$$

II°  
-  $AB \rightarrow D$

$$(A)_F^+ = AC \not\subset D \quad \text{Non TdG.}$$

$$(B)_F^+ = B \not\subset D \quad \text{Non TdG.}$$

-  $ABH \rightarrow C$

$$(A)_F^+ = AC \supset C \text{ TdG.} \quad \cancel{ABH \rightarrow C} \quad \text{MA GIÀ DEDUTTE}$$

$$F = \{A \rightarrow C, AB \rightarrow D, ABH \rightarrow E, BH \rightarrow A, BH \rightarrow D, D \rightarrow C, DE \rightarrow H, E \rightarrow B\}$$

$$-ABH \rightarrow E$$

$$(A)_F^+ = AC \quad \nexists E \quad \text{non TOLG.}$$

$$(B)_F^+ = B \quad \nexists E \quad \text{non TOLG.}$$

$$(H)_F^+ = H \quad \nexists E \quad \text{non TOLG.}$$

$$-BH \rightarrow A$$

$$(B)_F^+ = B \quad \nexists A$$

$$(H)_F^+ = H \quad \nexists A$$

$$-BH \rightarrow D \quad \text{non TOLG.}$$

$$-DE \rightarrow H$$

$$(D)_F^+ = DC \quad \nexists H$$

$$(E)_F^+ = EB \quad \nexists H$$

alle fnc del para  $\mathbb{I}^0$

$$F = \{ A \rightarrow C, AB \rightarrow D, ABH \rightarrow E, BH \rightarrow A, BH \rightarrow D, D \rightarrow C, DE \rightarrow H, E \rightarrow B \}$$

III<sup>a</sup> caso, idempotenza

$$- A \rightarrow C$$

$$(A)_{F-A \rightarrow C}^+ = A \quad \nexists C \quad \text{non TOLG.}$$

$$- AB \rightarrow D$$

$$(AB)_{F-AB \rightarrow D}^+ = ABC \quad \nexists D \quad \text{non TOLG.}$$

$$- ABH \rightarrow E$$

$$(ABD)^+ = ABD \cup D \cup E \quad \text{non tgl}$$

$$- BD \rightarrow D$$

$$(ABD)^+_{F=BD \rightarrow D} = BD \cup D \cup D \quad \text{tutto}$$

$$F = \{ A \rightarrow X, AB \rightarrow D, AB \rightarrow E, BD \rightarrow A, D \rightarrow C, DE \rightarrow H, E \rightarrow B \}$$

LE ALTRE LE VEDO A MEUTE

Alla fine del 3° Passo

$$F = \{ A \rightarrow X, AB \rightarrow D, AB \rightarrow E, BD \rightarrow A, D \rightarrow C, DE \rightarrow H, E \rightarrow B \}$$

$$P = \{ AC, AD, ABHE, BHA, DC, DEH, EB \}$$

$F'$  in 3NF PERCHÉ PRESENTA TUTTI GLI ELEMENTI DI  $R$

E HA IL SEN SENZA PERDITA POICHÉ CONTIENE ELEMENTI CHE  
CONTENGONO CHIAVI, ES. ABHE

## esercizio ③

3) E' dato un file di 168.600 record. Ogni record occupa 426 byte, di cui 32 per la chiave. Un blocco contiene 4096 byte. Un puntatore a blocco occupa 5 byte. Si utilizza una organizzazione ISAM

3a) Calcolare l'occupazione in blocchi del file principale considerando che tutti i blocchi contengono il massimo numero di record che consente di avere almeno il 20% di spazio libero

3b) Calcolare l'occupazione in blocchi dell'indice considerando i blocchi indice completamente pieni

3c) Calcolare il costo massimo di una ricerca

$$2a) \text{Block Size} = \left\lfloor \frac{\text{Block size} \cdot 80}{100} \right\rfloor = 3276$$

$$\text{Record} \times \text{Block} = \left\lceil \frac{\text{Block size}}{\text{Record size}} \right\rceil = \left\lceil \frac{3276}{426} \right\rceil = 7$$

$$3b) \text{Record} \times \text{Block Index} = \left\lceil \frac{\text{Block size}}{\text{Pointer size} + \text{key size}} \right\rceil = \left\lceil \frac{4096}{37} \right\rceil = 110$$

$$\text{Total Block Main} = \frac{\left\lceil \text{N}^{\circ} \text{Record} \right\rceil}{\text{Record} \times \text{Block}} = \frac{168600}{7} \approx 24086$$

$$\text{Total Block Index} = \frac{\left\lceil \text{Total Block Main} \right\rceil}{\text{Record} \times \text{Block Index}} = \frac{24086}{110} = 219$$

$$3c) \text{Avg Time } \lg_2 (\text{Total Block Index}) + 1 = 8$$