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)

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) $J := (\sigma_{CITTA' == MILANO \land N\text{-}PEZZI > 150}(MAGAZZINI \bowtie_{CODICE == CODICEM} SCORTE)) \bowtie_{CODICEP == CODICE} (\sigma_{PREZZO\_UNIT > 50}(PRODOTTI))$

Query finale: $Q = \pi_{PRODOTTI.CODICE,\, DESCRIZIONE,\, PREZZO\_UNIT,\, MAGAZZINI.CODICE,\, INDIRIZZO,\, N\text{-}PEZZI}(J)$

1b) $J = (PRODOTTI \bowtie_{CODICE == CODICEP} SCORTE) \bowtie_{CODICEM == CODICE} MAGAZZINI$ $\quad L2000 = \pi_{CODICEM}(\sigma_{NPEZZI \leq 2000}(J))$

$P500 = \pi_{CODICEM}(\sigma_{PREZZOUNIT \leq 500}(J))$ $\quad CODM = \pi_{CODICE}(MAGAZZINI)$

Query Finale: $Q = ((CODM - P500) \cup (CODM - L200)) \bowtie MAGAZZINI$

---

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

F = {A→C, AB→D, ABH→CE, BH→AD, D→C, DE→H, E→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)

$X_F^+ = \{Y \mid X \to Y \in F^+\}$

2a) Cercherò le X per cui $X_F^+ = R$, controllando poi che $\nexists X' \subset X \mid X'^+_F = R$!

C non sarà nella chiave, dato che non compare come determinante. Inizio con i sottoinsiemi di ABDEH, ma noto, osservando F, che $ABH \to CE \land ABH \to AB \land AB \to D \Rightarrow ABH_F^+ = R$, controllo i sottoinsiemi:

$AB^+ = ABCD$. $AH^+ = AHC$ $BH^+ = R$, controllo $B^+ = B$ e $H^+ = H$. $\Rightarrow$ BH e' chiave. Eviterò di controllare i sovrainsiemi di BH. Osservo però che $DE \to H \land D \to B$, sembra che DE faccia al caso nostro, infatti $DE_F^+ = R$, e $D^+ = DC \land E^+ = EB \Rightarrow$ DE e' chiave. Essendo che $AB \to D \in F$, saprò che $ABE_F^+ = R$ controllo $AE^+ = R$, $BE^+ = BE$, i sottoinsiemi di AE non sono chiave, quindi AE e' chiave. Osservo che $EH \to E \to B \land BH \to AD$, quindi calcolo $EH_F^+ = R \Rightarrow$ EH e' chiave. Le chiavi sono $\{AE, EH, DE, BH\}$.

2b) in $A \to C \in F$, A non e' superchiave, e C non e' primo, lo schema non e' in 3NF.

2c) Devo trovare una copertura minimale, inizio minimizzando i determinati:

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

Ora, cerco le $X \to Y \in F$ per cui $\exists X' \subset X \mid Y \in (X')^+_F$, cercando di minimizzare i determinanti.

- $AB \to D$: ho $A_F^+ = AC$ e $B^+ = B$. E' già minimale.
- $ABH \to C$: ho $A \to C \in F$, quindi posso eliminare $ABH \to C$ dato che $C \in A_F^+$.
- $ABH \to E$: ho che $BH_F^+ = R \Rightarrow$ sostituisco con $BH \to H$.
- $BH \to E \land BH \to D$ rimangono tali, dato che $B_F^+ = B \land H_F^+ = H$, analogo per $D \to C$ dato che $D_F^+ = D$
- $DE \to H$ e' già minimale dato che $E_F^+ = EB$.

La copertura minimale e': $F = \{A \to C, AB \to D, BH \to E, BH \to A, BH \to D, D \to C, DE \to H, E \to B\}$

```
ρ = ∅
S = ∅
for each A∈R | ∀x→y∈F, A∉xy {
        S = S∪A
}
if S≠∅ {ρ = ρ∪{S} ∧ R = R-S}
if ∃ x→y∈F | xy = R {
        ρ = ρ∪{R}
}
else {
        for each x→y∈F {
                ρ = ρ∪{xy}
        }
}
return ρ
```

Applico l'algoritmo. Non esiste $X \in R$ che non compaia in F, inoltre non esiste una dip. che comprenda tutto R, quindi $\rho = \{AC, ABD, BEH, ABH, BDH, CD, DEH, EB\}$, essendo che esiste $\{BEH\} \in \rho$ e $\underbrace{BH}_{\text{chiave}} \subseteq BEH$, ρ ha un Join senza perdita.

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

3a) devo considerare il 20% di un blocco, ossia $\lceil 4096 \cdot 0.2 \rceil = 820$ byte, quindi per ogni blocco, potrò usare al più $4096 - 820 = 3276$ byte, ogni blocco avrà $\lfloor \frac{3276}{426} \rfloor = 7$ record, necessitiamo in totale di $\lceil \frac{168\,600}{7} \rceil = 24086$ blocchi per il MainFile. 3b) In un blocco pieno, entrano $\lfloor \frac{4096}{32+5} \rfloor = 110$ coppie chiave puntatore, ne devo memorizzare 24086, necessito di $\lceil \frac{24086}{110} \rceil = 219$ blocchi.

3c) Il costo massimo di una ricerca binaria e' $\lceil \log_2(219) \rceil + 1 = 8 + 1 = 9$ accessi.

Il costo massimo per una ricerca sequenziale e' di 219 accessi.