

## Basi di Dati – Corso B – Appello: 10 dicembre 2018

Cognome, Nome \_\_\_\_\_

Matricola \_\_\_\_\_

### Domanda 1.

Si consideri la seguente base di dati (chiamata “Incarichi”) per la gestione degli incarichi da Primo Ministro della Repubblica Italiana:

POLITICO(Cognome, Nome, DataNascita)

ADESIONE(Cognome, Nome, NomePartito, DataAdesione)

PARTITO(NomePartito, Area\*, DataFondazione)

PRIMOMINISTRO(Cognome, Nome, DataIncarico, Durata\*)

Nella relazione PRIMOMINISTRO, *Cognome* e *Nome* costituiscono una chiave esterna su POLITICO. *DataIncarico* è la data in cui il politico avente *Cognome* e *Nome* riceve l’incarico (eventualmente esplorativo) da parte del Presidente della Repubblica. *Durata* contiene invece la eventuale durata (in giorni) del mandato ed è null se l’incarico non è andato a buon fine (cioè, se il politico non ha individuato una squadra di ministri e una maggioranza in parlamento).

In ADESIONE l’attributo *NomePartito* è una chiave esterna su PARTITO. *DataAdesione* indica invece la data in cui il politico *Cognome*, *Nome* ha aderito al partito politico indicato in *NomePartito*. Un politico **non può** appartenere a più partiti contemporaneamente.

In PARTITO l’attributo *Area* indica l’area di appartenenza del partito (destra, centro-destra, centro, centro-sinistra, sinistra). Se è null, il partito non fa riferimento a nessuna area politica tradizionale.

Si suppone che un Primo Ministro non possa cambiare partito politico durante il suo incarico.

Con riferimento alla base dati “Incarichi” si richiede di esprimere in SQL la seguente richiesta:

**Individuare l’area politica tradizionale che ha avuto il maggior numero di politici incaricati che hanno effettivamente governato, cioè i cui incarichi sono andati a buon fine.**

### Soluzione 1.

Una possibile soluzione è la seguente:

```
with PrimoMinSuaArea as (  
    select distinct(p.Cognome, p.Nome, p.Area)  
    from PRIMOMINISTRO pm join POLITICO p using (Cognome, Nome)  
    join ADESIONE a using (Cognome, Nome)  
    join PARTITO pp using (NomePartito)  
    where (pm.Durata is not null)  
    and (pp.Area is not null)  
    and (pm.DataIncarico>=a.DataAdesione)  
    and NOT EXISTS (select *  
                     from ADESIONE a2  
                     where (a2.DataAdesione>a.DataAdesione)  
                           and (a2.DataAdesione<=pm.DataIncarico))  
),  
with AreaQuantipM as (  
    select Area, count(*) as NroIncSucc  
    from PrimoMinSuaArea
```

```

        group by Area
    ),
    select Area, NroIncSucc
    from AreaQuantipM
    where NroIncSucc = (select max(NroIncSucc)
                        from AreaQuantipM);

```

### Domanda 2.

Con riferimento alla base dati “Incarichi” della Domanda 1, esprimere, in **algebra relazionale** e in **calcolo relazionale** su tuple con dichiarazione di range, l'interrogazione:

**Indicare cognome, nome e data di nascita dei politici che hanno aderito a partiti di area diversa.**

### Soluzione 2.

Una possibile soluzione in algebra relazionale è la seguente:

$$\begin{aligned}
 & \text{POLITICO} \bowtie_{\Theta_1} \\
 & (\rho_{A1 \leftarrow \text{ADESIONE}}(\text{ADESIONE}) \bowtie_{\Theta_2} \rho_{PP1 \leftarrow \text{PARTITOPOLITICO}}(\text{PARTITO})) \\
 & \quad \bowtie_{\Theta_3} \\
 & (\rho_{A2 \leftarrow \text{ADESIONE}}(\text{ADESIONE}) \bowtie_{\Theta_4} \rho_{PP2 \leftarrow \text{PARTITOPOLITICO}}(\text{PARTITO}))
 \end{aligned}$$

Dove:

- $\Theta_1$ :  $\text{POLITICO.Cognome} = A1.Cognome \wedge \text{POLITICO.Nome} = A1.Nome$
- $\Theta_2$ :  $A1.NomePartito = PP1.NomePartito$
- $\Theta_3$ :  $(PP1.Area \neq PP2.Area) \vee (PP1.Area \text{ IS NULL} \wedge PP2.Area \text{ IS NOT NULL})$   
 $\vee (PP1.Area \text{ IS NOT NULL} \wedge PP2.Area \text{ IS NULL}) \wedge$   
 $(A1.Cognome = A2.Cognome \wedge A1.Nome = A2.Nome)$
- $\Theta_4$ :  $A2.PartitoPolitico = PP2.PartitoPolitico$

Una possibile soluzione in calcolo relazionale è la seguente:

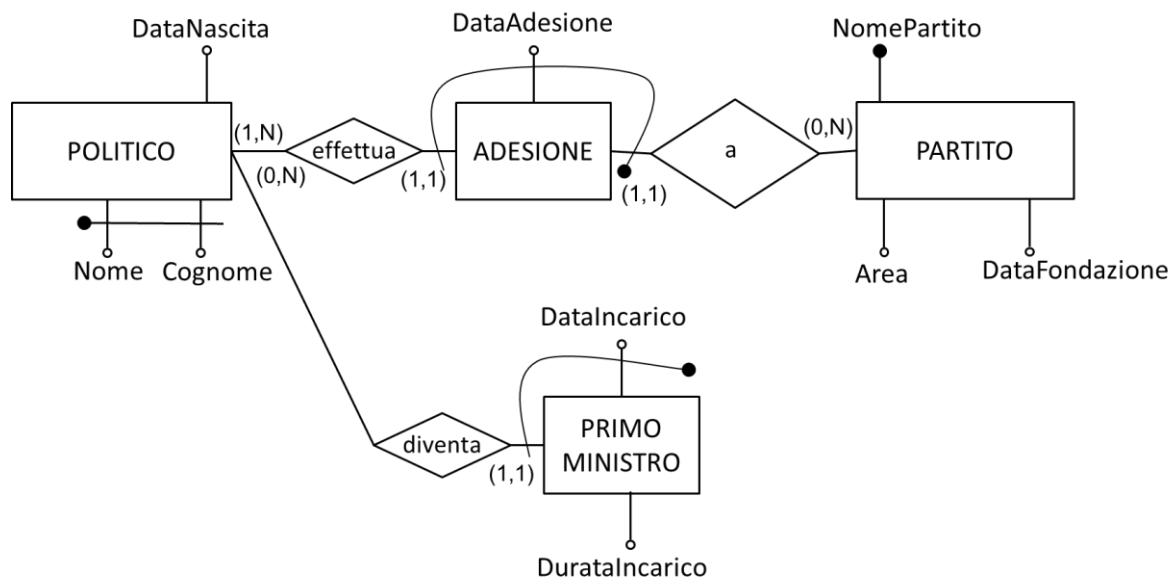
$$\begin{aligned}
 & \{ p.* \mid p(\text{POLITICO}) \mid \exists a(\text{ADESIONE}) \exists pp(\text{PARTITO}) \exists a'(\text{ADESIONE}) \exists pp'(\text{PARTITO}) \\
 & \quad (a.Cognome = p.Cognome \wedge a.Nome = p.Nome \wedge a.NomePartito = pp.NomePartito) \wedge \\
 & \quad (a'.Cognome = a.Cognome \wedge a'.Nome = a.Nome \wedge a'.NomePartito = pp'.NomePartito) \wedge \\
 & \quad (pp.Area \neq pp'.Area) \vee (pp.Area \text{ IS NULL} \wedge pp'.Area \text{ IS NOT NULL}) \\
 & \quad \vee (pp.Area \text{ IS NOT NULL} \wedge pp'.Area \text{ IS NULL}) \}
 \end{aligned}$$

### Domanda 3.

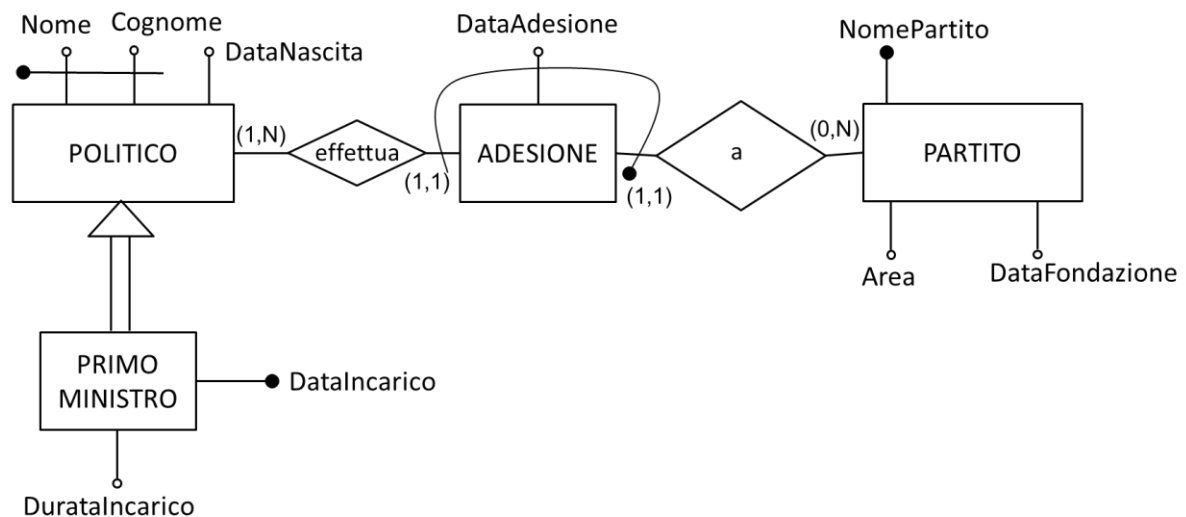
Progettare uno schema concettuale ER che specifichi la realtà considerata nella base dati "Incarichi" della Domanda 1. Elencare, sotto forma di regole aziendali, tutti i vincoli non esprimibili attraverso l'ER.

### Soluzione 3.

Una possibile soluzione è la seguente:



oppure



con i seguenti vincoli e regole aziendali:

- V1.  $PARTITO.Area \in \{destra, centro-destra, centro, centro-sinistra, sinistra\}$
- R1. Un politico, nel periodo in cui è in carica come primo ministro, non può cambiare partito.
- R2. Un politico non può aderire a più partiti contemporaneamente.
- R3. Se  $PRIMOMINISTRO.Durata$  IS NULL, l'incarico non è andato a buon fine
- R4. Se  $PARTITO.Area$  IS NULL, il partito non appartiene a nessun'area politica tradizionale.

#### Domanda 4.

Si consideri la relazione ERBARIO relativa alla classificazione in un erbario di fiori vari che crescono in Italia col loro periodo di fioritura. Chiave di ERBARIO è la coppia {Codice, Provincia} mentre F è l'insieme di dipendenze funzionali. Dire se ERBARIO è in 3FN e se non lo è scomporla.

ERBARIO(Codice, Provincia, NomeScientifico, DataClassif, Scopritore, PeriodoFiore, LinkScopritore)

con  $F = \{ \text{Codice} \rightarrow \text{NomeScientifico}, \text{DataClassif}, \text{Scopritore}, \text{PeriodoFiore};$

$\text{Scopritore} \rightarrow \text{LinkScopritore};$

$\text{NomeScientifico} \rightarrow \text{Codice}, \text{DataClassif}, \text{Scopritore}, \text{PeriodoFiore} \}$

#### Soluzione 4.

La relazione ERBARIO non è in 3FN, in quanto nessuna delle dipendenze funzionali di F è in 3NF.

Per procedere alla sintesi in 3FN, bisogna prima calcolare la copertura minimale di F. A tal fine si deve prima porre F in forma canonica:

$$F = \begin{cases} F1: \text{Codice} \rightarrow \text{NomeScientifico}; \\ F2: \text{Codice} \rightarrow \text{DataClassif}; \\ F3: \text{Codice} \rightarrow \text{Scopritore}; \\ F4: \text{Codice} \rightarrow \text{PeriodoFiore}; \\ F5: \text{Scopritore} \rightarrow \text{LinkScopritore}; \\ F6: \text{NomeScientifico} \rightarrow \text{Codice}; \\ F7: \text{NomeScientifico} \rightarrow \text{DataClassif}; \\ F8: \text{NomeScientifico} \rightarrow \text{Scopritore}; \\ F9: \text{NomeScientifico} \rightarrow \text{PeriodoFiore} \end{cases}$$

Ovviamente non esistono attributi estranei. Esistono però dipendenze ridondanti. A seconda dell'attributo ponte considerato (Codice o NomeScientifico) si ottengono soluzioni diverse. Considerando Codice, ad esempio, F7, F8 ed F9 diventano ridondanti per transitività. L'insieme di copertura minimale diventa quindi:

$$F' = \begin{cases} F1: \text{Codice} \rightarrow \text{NomeScientifico}; \\ F2: \text{Codice} \rightarrow \text{DataClassif}; \\ F3: \text{Codice} \rightarrow \text{Scopritore}; \\ F4: \text{Codice} \rightarrow \text{PeriodoFiore}; \\ F5: \text{Scopritore} \rightarrow \text{LinkScopritore}; \\ F6: \text{NomeScientifico} \rightarrow \text{Codice}; \end{cases}$$

e accorpendo tutte le d.f. con lo stesso antecedente:

$$F' = \begin{cases} F1: \text{Codice} \rightarrow \text{NomeScientifico, DataClassif, Scopritore, PeriodoFiore}; \\ F2: \text{Scopritore} \rightarrow \text{LinkScopritore}; \\ F3: \text{NomeScientifico} \rightarrow \text{Codice}; \end{cases}$$

Si può quindi procedere alla scomposizione, ottenendo le seguenti relazioni in 3FN:

R1(Codice, NomeScientifico, DataClassif, Scopritore, PeriodoFiore)

R2(Scopritore, LinkScopritore)

R3(NomeScientifico, Codice)

cui bisogna aggiungere la relazione contenente la chiave (Provincia,Codice):

R4(Provincia, Codice)

Infine, è possibile accorpare R1 ed R3, in quanto R3 è contenuta in R1. La soluzione è quindi:

R1(Codice, NomeScientifico, DataClassif, Scopritore, PeriodoFiore)

con F1={ NomeScientifico  $\rightarrow$  Codice }

R2(Scopritore, LinkScopritore)

R4(Provincia, Codice)

#### **Domanda 5.**

- A. Esporre la differenza tra indici densi ed indici sparsi.
- B. Indicare almeno **due casi** in cui gli indici secondari si dimostrano inefficienti.

#### **Soluzione 5.**

Si vedano gli appunti/slide/testo.