

Esercizio: Esame

In questo esercizio sono dettagliati tutti i passi richiesti per la prima consegna della tesina

Il punto di partenza è il diagramma relazionale del DBO con la relativa documentazione

Per ognuno dei FATTI, i passi principali da svolgere sono

A) DATA PROFILING

B) SCHEMA RELAZIONALE

C) SCHEMA ER

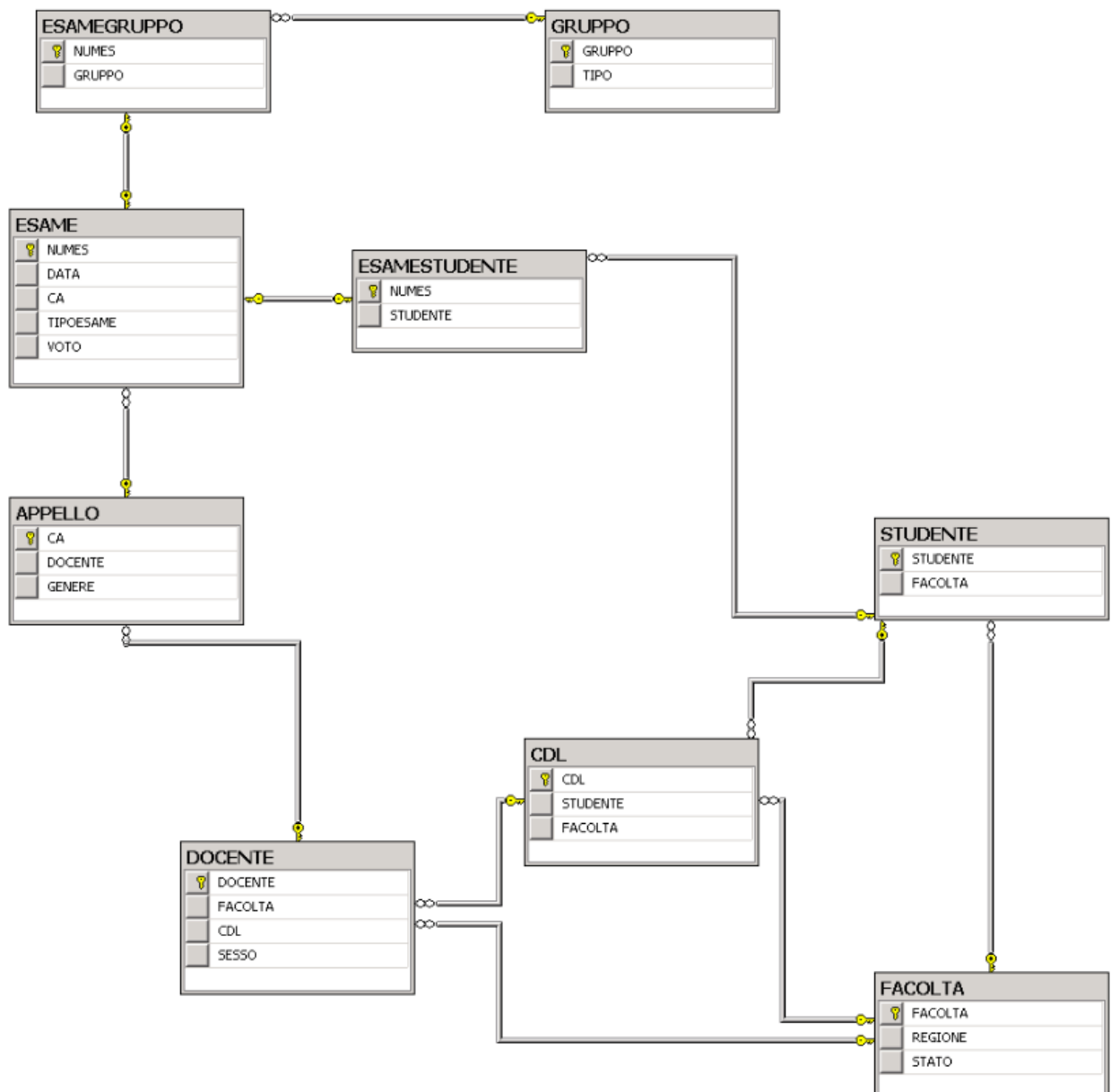
D) PROGETTO CONCETTUALE

- a. ALBERO DEGLI ATTRIBUTI (per chi usa indico, parta da un DFM completo di tutti gli attributi in quanto indyco permette di disegnare solo DFM)
- b. SCHEMA DI FATTO (il DFM)

Questo Esempio è stato iniziato in laboratorio il 22 Ottobre

Il database utilizzato è molto simile a quello disponibile in questo BackUp
<http://www.dbgroup.unimo.it/SIA/OT22.BAK>

DIAGRAMMA RELAZIONALE



Documentazione:

- FK: CDL → FACOLTA : un CDL (CorsoDiLaurea) ha sede in una FACOLTA
- FK: DOCENTE → CDL : un DOCENTE è di un CDL
- FK: APPELLO → DOCENTE: un APPELLO è con il DOCENTE che lo tiene
- FK: CDL → STUDENTE: un CDL ha uno STUDENTE rappresentante

Attributo **TIPO_ESAME** di ESAME : assume il valore STUD se è un esame di uno studente ed il valore GRUP se è un esame di un gruppo

CDL è sinonimo di CDS

A) DATA PROFILING:

Chiavi Alternative:

in CDS , ho AK: STUDENTE ? Si effettuano le query relative

```
SELECT *  
FROM CDL  
WHERE STUDENTE IS NULL
```

```
SELECT STUDENTE  
FROM CDL  
GROUP BY STUDENTE  
HAVING COUNT(*) > 1
```

La seconda query ha un risultato non vuoto quindi la risposta è negativa, cioè non ho AK quindi uno studente può essere rappresentante in più CDL!!

Dipendenza funzionale:

in FACOLTA , ho FD: REGIONE → STATO ? Si effettua la query relativa

```
SELECT REGIONE  
FROM FACOLTA  
GROUP BY REGIONE  
HAVING COUNT(DISTINCT STATO) > 1
```

e la risposta è positiva in quanto la query non restituisce niente.

A) SCHEMA RELAZIONALE

Viene riportato lo schema relazionale, comprensivo di key (primary Key ed eventuali Alternative Key), Foreign Key e Dipendenze Funzionali

FACOLTA (FACOLTA, REGIONE, STATO)

FD: REGIONE → STATO

CDS (CDS, FACOLTA:FACOLTA, STUDENTE:STUDENTE)

DOCENTE (DOCENTE, FACOLTA:FACOLTA, CDL:CDL, SESSO)

STUDENTE (STUDENTE, FACOLTA:FACOLTA)

APPELLO (APPELLO, DOCENTE:DOCENTE, GENERE)

ESAME (NUMES, APPELLO:APPELLO, DATA, TIPO_ESAME, VOTO)

ESAMEGRUPPO (NUMES:ESAME, GRUPPO:GRUPPO)

ESAMESTUDENTE (NUMES:ESAME, STUDENTE:STUDENTE)

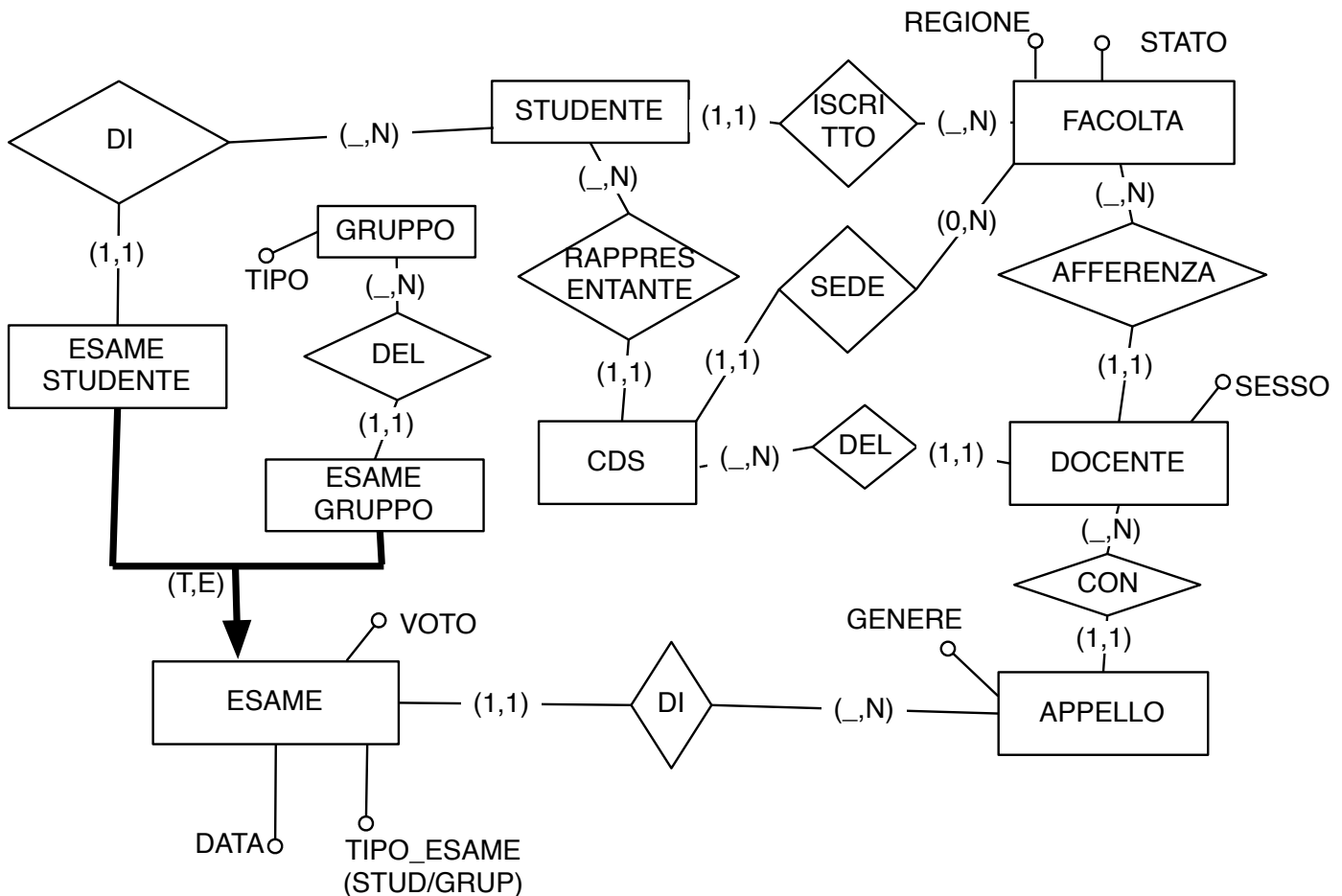
GRUPPO (GRUPPO, TIPO)

C) SCHEMA ER

Effettuare il reverse-engineering ed ottenere il relativo schema ER.

Vedere http://www.dbgroup.unimo.it/SIA/EsempioReverseEngineeringSIA_2014.pdf e

http://www.dbgroup.unimo.it/SIA/EsempioPerTesinaOttobre2014_Soluzione.pdf



NOTE

- 1) Le chiavi primarie, cioè gli identificatori primari delle entità, non sono indicati, in quanto *sottintesi*: ogni entità ha come identificatore il singolo attributo che costituisce la chiave primaria
- 2) Sono stati messi anche i nomi alle associazioni binarie uno-a-molti: i nomi non sono *importanti*, servono per indicare l'*interpretazione* del progettista (un **DOCENTE** è di una **FACOLTA** viene specificato dal progettista come **AFFERENZA**, cioè un **DOCENTE** afferisce ad una **FACOLTA**)
- 3) Le cardinalità minime nel caso (1,N) sono lasciate indefinite; nella tesina occorre riportarle (vedere esempio http://www.dbgroup.unimo.it/SIA/EsempioReverseEngineeringSIA_2014.pdf e
- 4) Si suppone sia stato verificato sui dati che la gerarchia su **ESAME** sia totale/esclusiva (simile al controllo fatto in aula sulla gerarchia su **BusinessEntity** ...)

D) PROGETTO CONCETTUALE

FATTO: ESAME

DIMENSIONI: GRUPPO, STUDENTE, TIPOESAME, DATA, DOCENTE

ANALISI:

Si vogliono **analizzare gli esami** senza scendere nel dettaglio del singolo APPELLO ma considerando solo il DOCENTE che ha tenuto l'APPELLO. Si vuole effettuare l'analisi anche rispetto allo studente e/o gruppo che ha sostenuto l'esame, al tipo di esame e alla data.

MISURE:

per la prima consegna limitarsi a definire le misure indicandone il nome e se eventualmente prevedete che siano delle medie, dei MAX, etc. Non servono per la prima consegna altri dettagli sulle misure. In questo esempio:

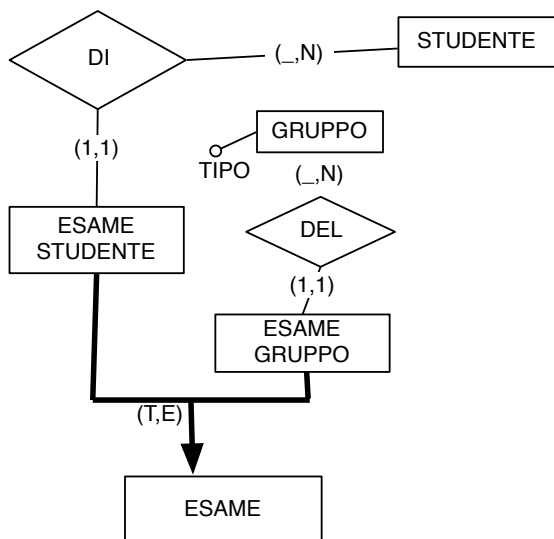
NUMERO ESAMI (SOMMA), cioè il numero degli esami fatti

VOTO ESAME (MEDIA), cioè il voto riportato all'esame

E' importante osservare che una misura, ad esempio NUMERO ESAMI verrà poi calcolata rispetto alle dimensioni (in base al suo operatore di aggregazione, in questo caso la SUM, cioè misura addittiva): pertanto è sbagliato riportare anche altre misure quali NUMERO ESAMI GRUPPO intendendo il numero di esami che un gruppo ha fatto, in quanto questa è la misura NUMERO ESAMI valutata per un certo gruppo

Cominciamo la progettazione:

ESAME ha una gerarchia composta da ESAME_STUDENTE (con lo STUDENTE ASSOCIATO) ed ESAME_GRUPPO (con il GRUPPO ASSOCIATO)



Come spiegato nell'esempio di pag. 25 delle dispense

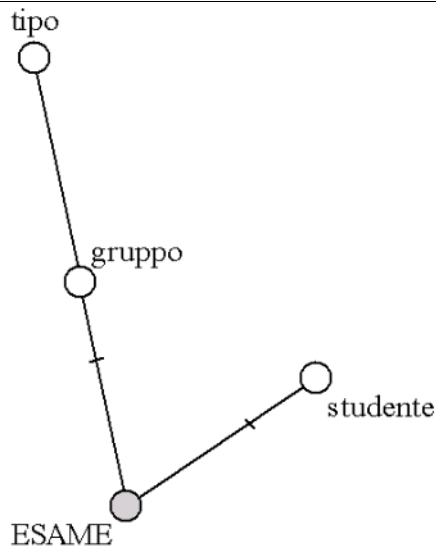
http://www.dbgroup.unimo.it/SIA/SIA_2015_IIModelloConcettualeDFM.pdf

ed in maniera analoga a quanto fatto nell'esempio del BIGLIETTO con CHECK_IN, considerando come radice ESAME, la gerarchia viene rappresentata nell'albero degli attributi nel seguente modo

Quando un ESAME è un ESAME_STUDENTE avrà associato uno STUDENTE (che risulta quindi opzionale)

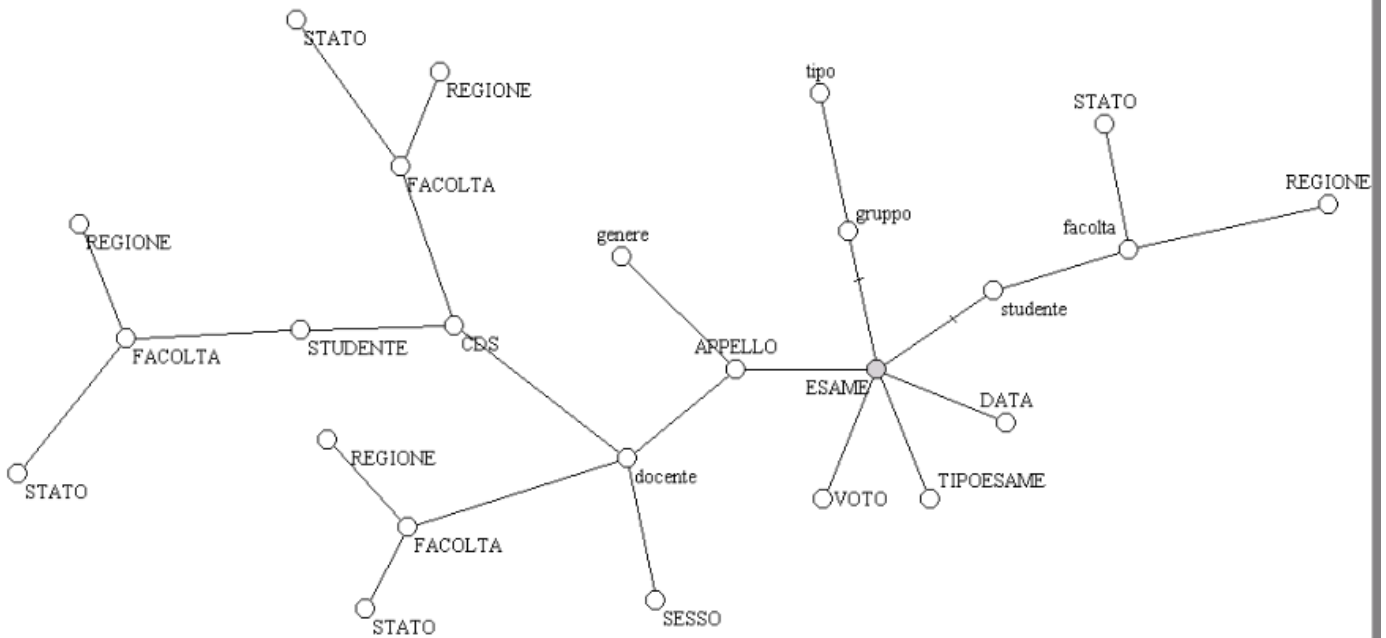
Quando un ESAME è un ESAME_GRUPPO avrà associato uno GRUPPO (che risulta quindi opzionale); al gruppo è associato il TIPO

Si può aggiungere nell'albero degli attributi (oppure nella documentazione allegata) il fatto che GRUPPO e STUDENTE rispetto ad ESAME sono Totale ed Esclusivo. Questo è un dettaglio che verrà considerato nel prossimo esercizio.



A questo punto si costruisce il resto dell'albero degli attributi in modo **praticamente automatico**:

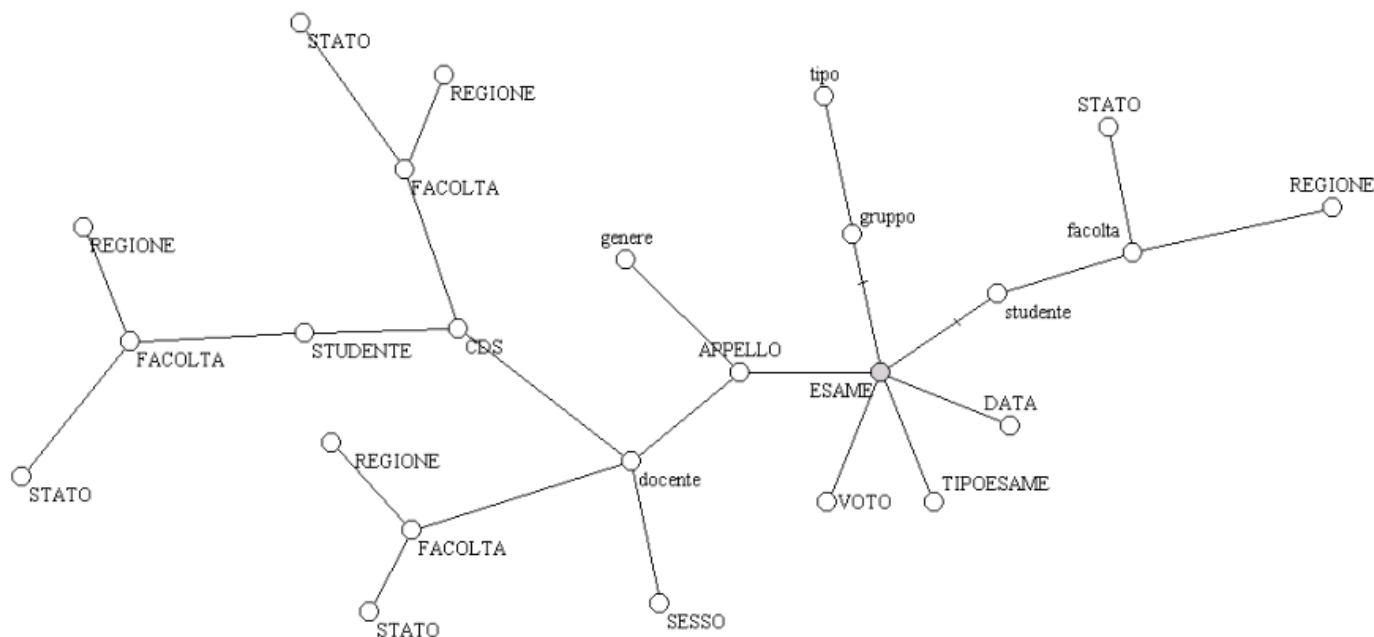
- per ogni associazione uno-a-molti si aggiunge il relativo ramo e si continua a sviluppare l'albero
- per ogni attributo si aggiunge il relativo ramo e ci si ferma



USO DI Indyco (<http://www.indycosuite.com/>)

Da quest'anno la parte di **Editing dell'Albero** e di disegno dello **Schema di Fatto** può essere effettuata tramite Indyco: l'uso di tale strumento è facoltativo ma fortemente consigliato per semplificare - rispetto ad altri editor classici quali PPT - il disegno dell'albero (e quindi delle gerarchie nelle dimensioni) e dello Schema di Fatto.

In caso di uso indyco, l'albero degli attributi da cui partire dovrà essere disegnato come uno schema di fatto (Indyco infatti non permette di scotriure alberi degli attributi ma solo DFM). Quindi chi usa Indico effettuerà la parte di tesina legata all'albero degli attributi sostituendo nodo (il pallino) ESAME con il fatto ESAME. Il resto della trattazione rimane la stessa.



Una volta disegnato l'albero iniziale o il DFM iniziale per chi usa indyco, si può procedere **all'editing**, ovvero per effettuare potature/innesti, aggiunta di dipendenze funzionali, ottenendo quindi lo Schema di Fatto finale ESAME

IMPORTANTE

Nella progettazione (e quindi nell'albero) riportare solo gli attributi *significativi* dello schema relazionale: con riferimento alla tesina, attributi quali **rowguid**, **Modified Date** o altri attributi che avete già deciso di non usare (ad esempio attributi che non hanno valore nel DB) non dovrebbero essere utilizzati

L'albero degli attributi evidenzia che nel nostro database per gli esami sono riportate **quattro facoltà**

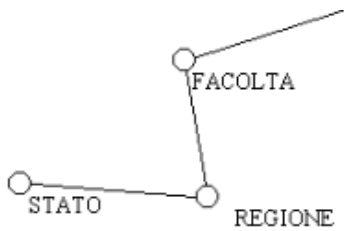
- 1) la facoltà del docente dell'appello dell'esame
- 2) la facoltà del CDS del docente dell'appello dell'esame
- 3) la facoltà dello studente rappresentante del CDS
- 4) la facoltà dello studente (eventuale) che ha sostenuto l'esame

E' ovviamente un caso limite, però non sono rari i casi in cui uno stesso concetto (la FACOLTA nel nostro esempio) compare due o più volte nell'analisi di un dato fatto.

Inizialmente si considera il caso generale di CONDIVISIONE: queste quattro facoltà possono essere quattro facoltà distinte. Poi successivamente si effettua l'analisi di eventuali convergenze, cioè si analizza se alcune di queste facoltà sono coincidenti

Editing dell'albero: SI AGGIUNGONO LE FD individuate (e non riportate nello schema E/R). Nel nostro esempio la FD in FACOLTA : REGIONE → STATO

Grazie alla condivisione questa operazione viene fatta una sola volta



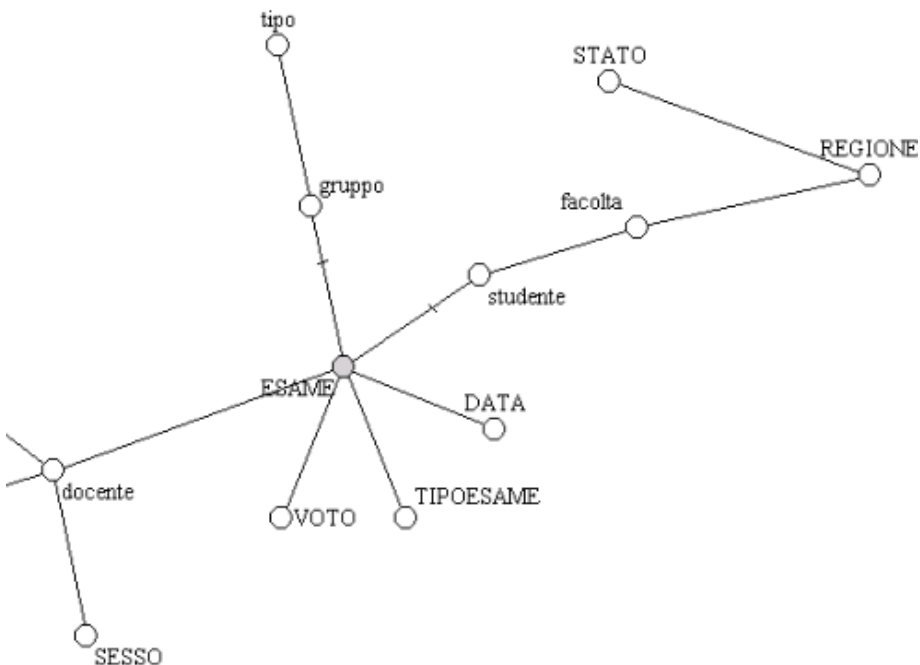
Consideriamo ora le dimensioni

GRUPPO, STUDENTE, TIPOESAME, DATA, DOCENTE

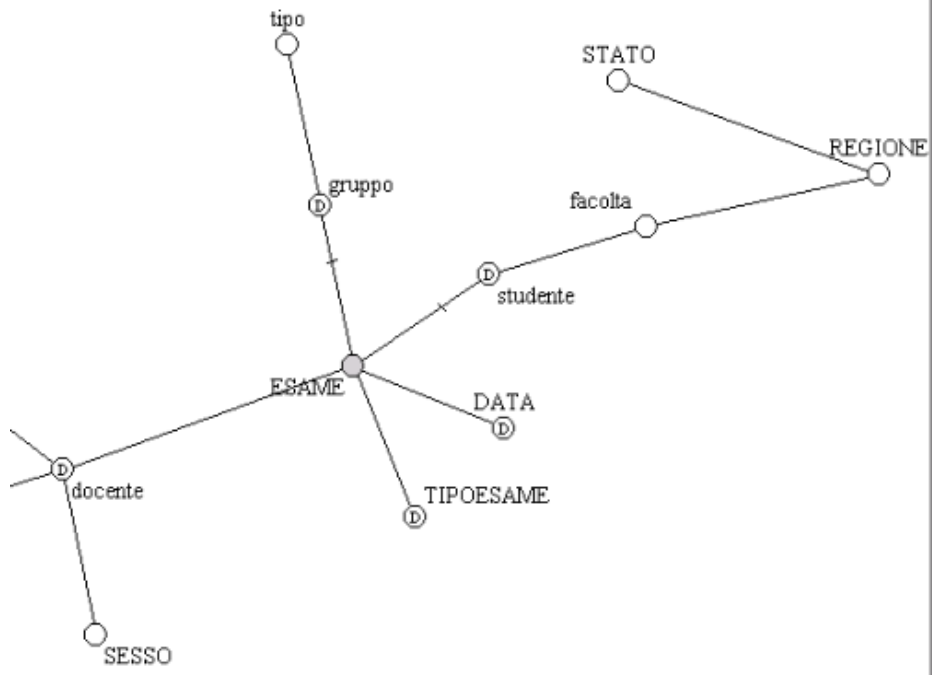
Le dimensioni **GRUPPO, STUDENTE, TIPOESAME** sono figli diretti della radice e quindi possono essere presi come dimensioni senza ulteriori elaborazioni; in particolare le due dimensioni **GRUPPO e STUDENTE** risultano opzionali

La dimensione **DOCENTE** non è figlio diretto della radice: questo significa che devo *togliere* qualcosa, cioè l'APPELLO: devo fare un innesto su APPELLO, portare **DOCENTE** come figlio della radice (il **GENERE** quindi scompare, non mi interessa)

La parte di albero modificata risulterà quindi:



VOTO viene considerata come MISURA e quindi viene *tolta* ; vengono indicate le dimensioni

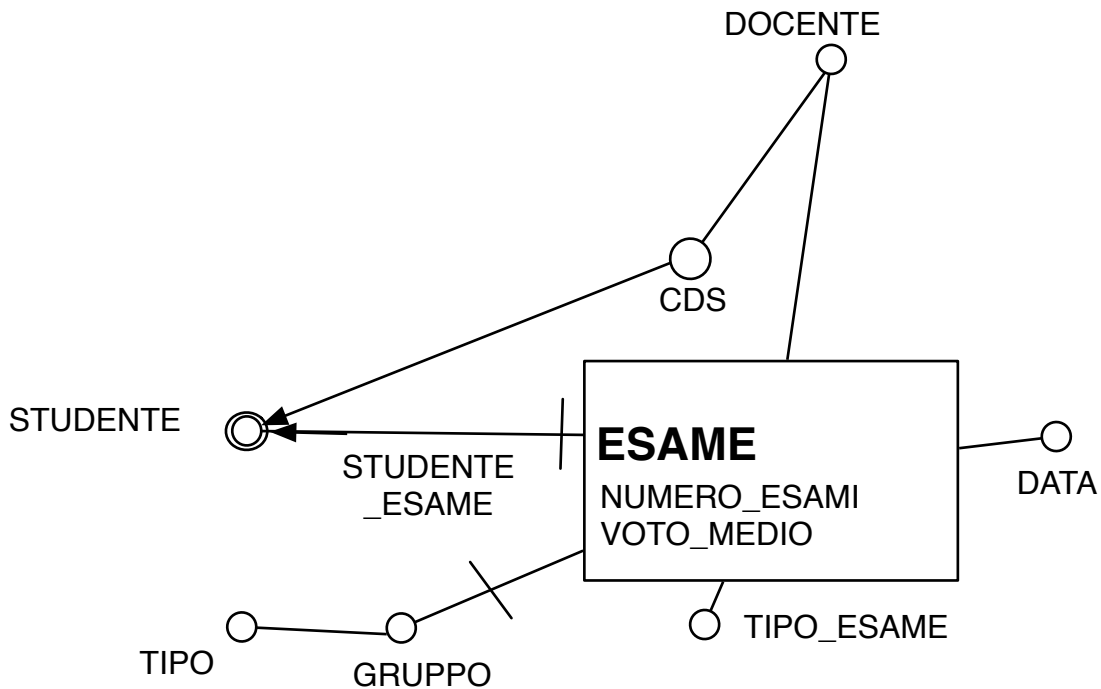


Schema di Fatto ESAME

Si costruisce lo Schema di Fatto ESAME con le quattro dimensioni; vengono indicate anche le misure

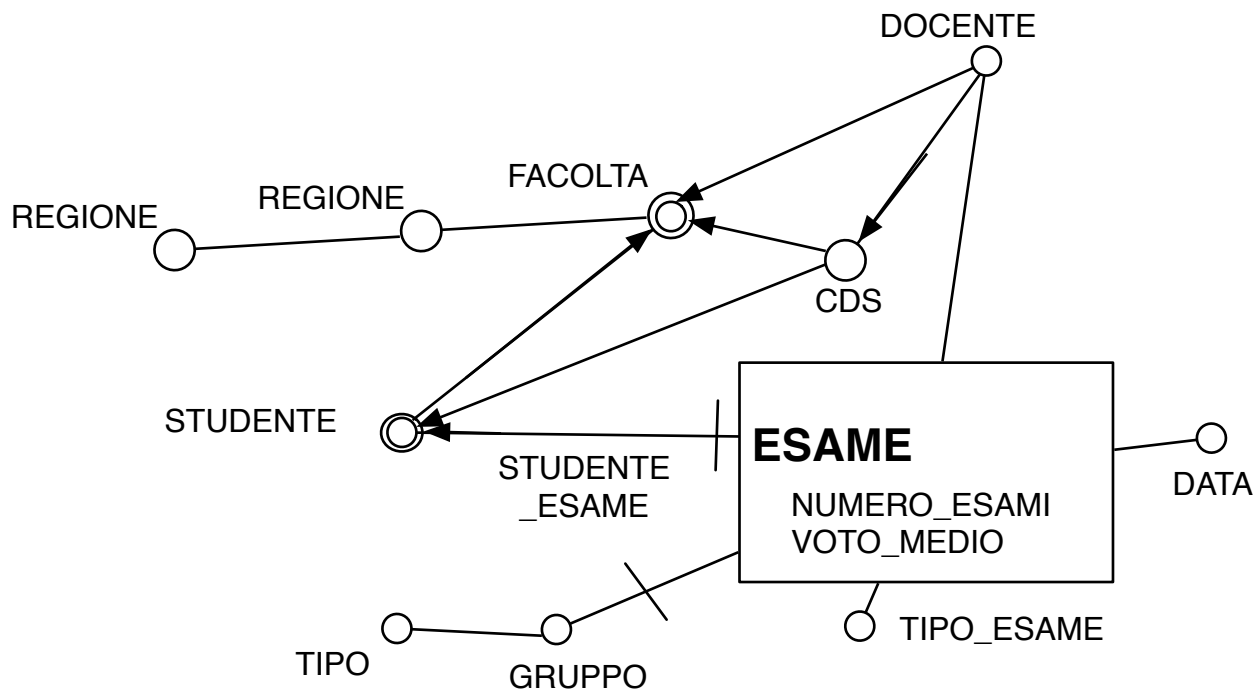
NUMERO_ESAMI, cioè il numero degli esami fatti
VOTO_ESAME_MEDIO, cioè il voto medio riportato all'esame

Per semplicità riportiamo prima la parte senza FACOLTA



Si noti condivisione su **STUDENTE** e la necessità di dare un nome al ruolo **STUDENTE_ESAME** : è appunto la dimensione **STUDENTE** richiesta nelle specifiche (in altre parole, visto che **STUDENTE** è condiviso la dimensione richiesta "STUDENTE dell'esame " è stata chiamata **STUDENTE_ESAME**

Ora aggiungiamo FACOLTA ottenendo lo **Schema di Fatto** finale per **ESAME**



Controlliamo che ci siano tutte e quattro le facoltà richieste

- 1) la facoltà del docente dell'appello dell'esame
DOCENTE.FACOLTA
- 2) la facoltà del CDS del docente dell'appello dell'esame
DOCENTE.CDS.FACOLTA
- 3) la facoltà dello studente rappresentante del CDS
DOCENTE.CDS.STUDENTE.FACOLTA
- 4) la facoltà dello studente (eventuale) che ha sostenuto l'esame
STUDENTE_ESAME.FACOLTA

SCHEMA DI FATTO **TEMPORALE/TRANSAZIONALE**

E' un argomento che verrà svolto la prossima settimana

Il controllo è molto semplice

L'insieme delle dimensioni

{ GRUPPO, STUDENTE, TIPOESAME, DATA, DOCENTE }

contiene **almeno una** delle chiavi della relazione dello schema di partenza che ho scelto come fatto, cioè della relazione ESAME?

Se la risposta è negativa (come in questo caso) allora lo schema di fatto è **TEMPORALE**
Altrimenti è **TRANSAZIONALE**

Quello svolto finora è la **parte minima indispensabile per la prima consegna**

Le misure verranno considerate nella **seconda consegna**

I seguenti punti **devono essere discussi nella tesina**

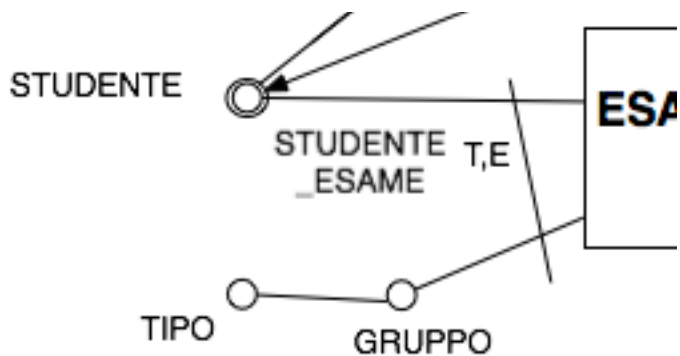
- 1) **COSTRUTTI AVANZATI**, quali attributi cross-dimensionali, archi opzionali e coperture
- 2) **ANALISI CONVERGENZE**
- 3) **DIPENDENZE FUNZIONALI** tra le DIMENSIONI

Essi possono essere inclusi indifferentemente nella **prima consegna oppure nella seconda consegna**

COSTRUTTI AVANZATI

Copertura di Arco Opzionale

Si conclude lo schema di fatto indicando alcuni *dettagli* (costrutti avanzati); ad esempio in questo caso possiamo riportare che tra dimensione GRUPPO e la dimensione STUDENTE_ESAME c'è una **Copertura di Arco Opzionale** di tipo Totale Esclusiva: cioè le due dimensioni sono opzionali, non posso avere sia lo STUDENTE_ESAME che il gruppo (Esclusiva) ed inoltre un ESAME avrà necessariamente o uno STUDENTE_ESAME o un GRUPPO



ANALISI CONVERGENZE

Un altro aspetto importante da discutere sullo schema di fatto sono le **eventuali convergenze**. Cioè si devono considerare le varie condivisioni individuate e si deve discutere se sono invece delle convergenze.

- 1) Ci sono dei casi in cui la risposta è immediata: la condivisione su STUDENTE non è (non ha senso che sia) una convergenza, altrimenti andrei a considerare solo gli esami degli studenti che sono rappresentati nel CDL del docente che ha fatto l'esame ...
- 2) In altri casi si devono fare delle indagini sulla disposizione, come nel caso di FACOLTA.

Si considerano le coppie di FACOLTA. Controlliamo che ci siano tutte e quattro le facoltà richieste.

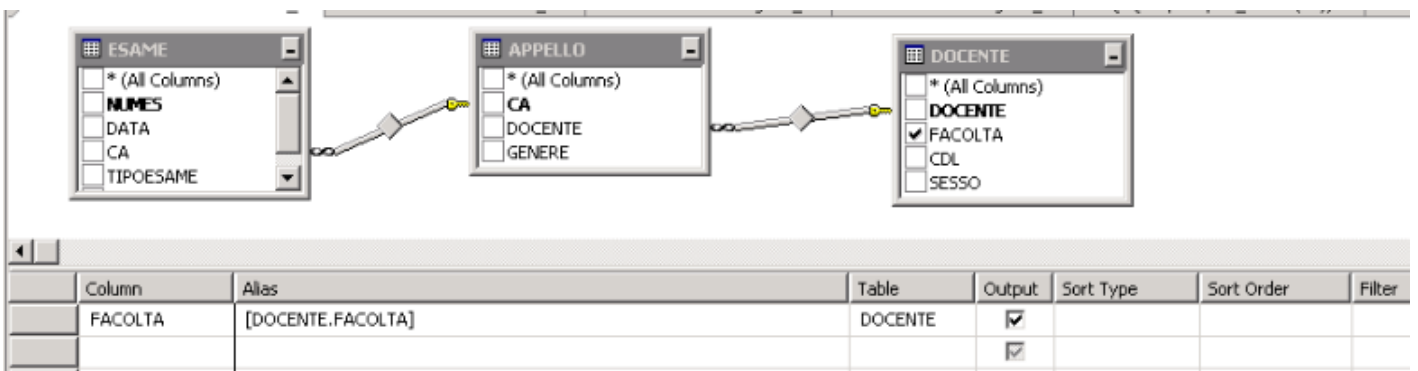
- 1) la facoltà del docente dell'appello dell'esame DOCENTE.FACOLTA
- 2) la facoltà del CDS del docente dell'appello dell'esame DOCENTE.CDS.FACOLTA

Convergono ??

Effettuo il controllo sui dati tramite una query che mi restituisce queste due facoltà: Sarà una query con vari join quindi la costruisco tramite l'interfaccia delle viste.

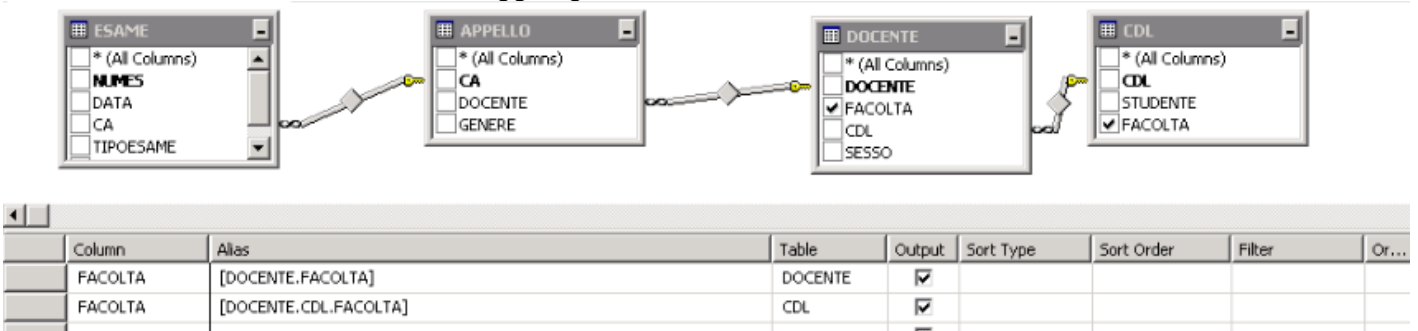
Guardo il Diagramma relazionale per capire dove sono queste due FACOLTA.

- 1) la facoltà del docente dell'appello dell'esame DOCENTE.FACOLTA
è in DOCENTE associato all'APPELLO dell'ESAME



Quindi seleziono FACOLTA e gli assegno alias DOCENTE.FACOLTA.

- 2) la facoltà del CDS del docente dell'appello dell'esame DOCENTE.CDS.FACOLTA:
è in CDL (ricordiamo che CDS è sinonimo di CDL) del DOCENTE associato all'APPELLO dell'ESAME : è importante notare che è sempre lo stesso DOCENTE, quindi non devo aggiungere un'altra volta la tabella DOCENTE, ma devo aggiungere solo la tabella CDL.



Quindi seleziono FACOLTA in CDL e gli assegno alias DOCENTE.CDL.FACOLTA.

Eseguo query e vedo che sono diverse

| | DOCENTE.FACOLTA | DOCENTE.CDL.FACOLTA |
|--|-----------------|---------------------|
| | MUSICA | LETTERE |
| | LETTERE | MUSICA |
| | LETTERE | MUSICA |
| | MUSICA | LETTERE |
| | LETTERE | MUSICA |
| | LETTERE | MUSICA |

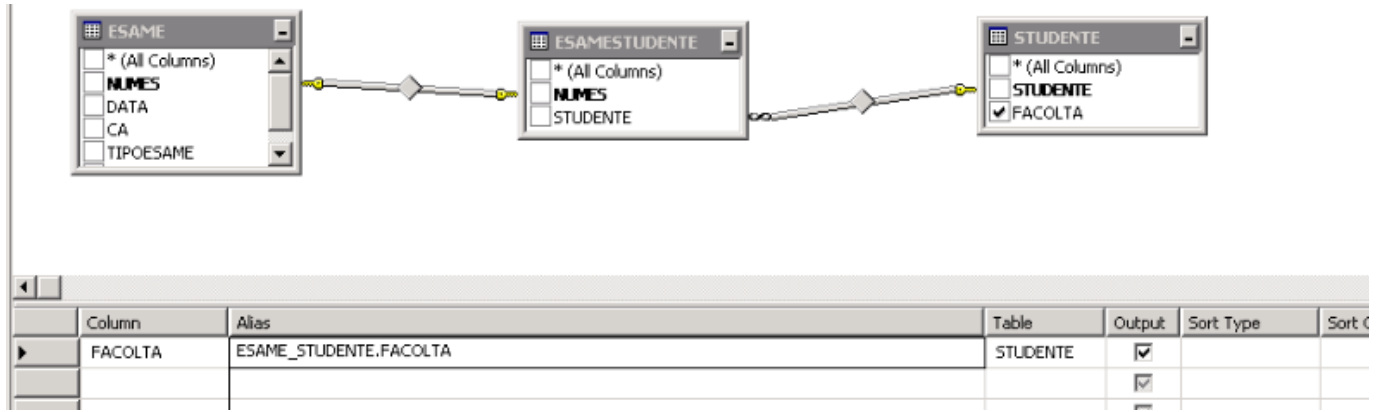
Per fare il controllo effettivo, alla query generata aggiungo la condizione WHERE

WHERE dbo.DOCENTE.FACOLTA <> dbo.CDL.FACOLTA

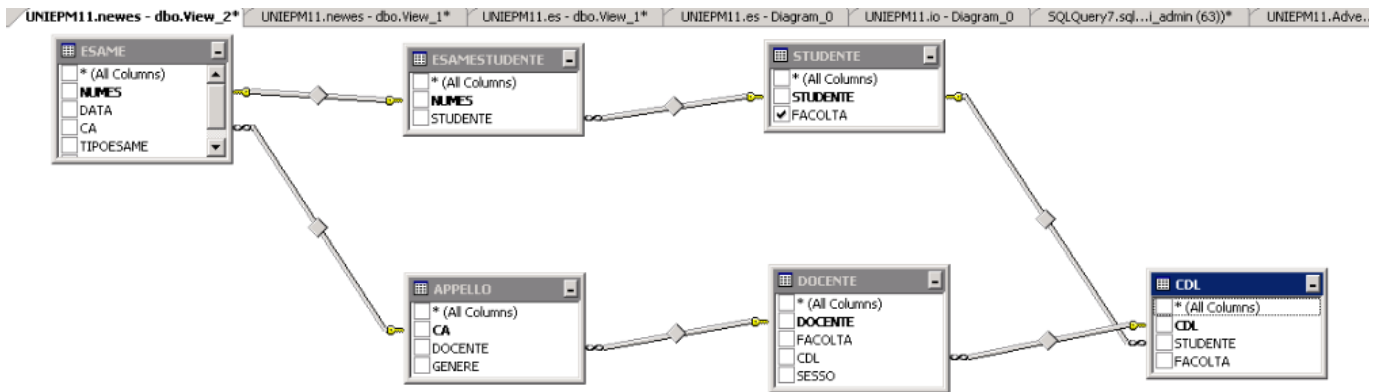
Eseguo la query: se il **risultato è vuoto** allora c'è **convergenza**, altrimenti resta condivisione

Stessa cosa per le altre coppie (questo è un caso **molto, molto** particolare ...)

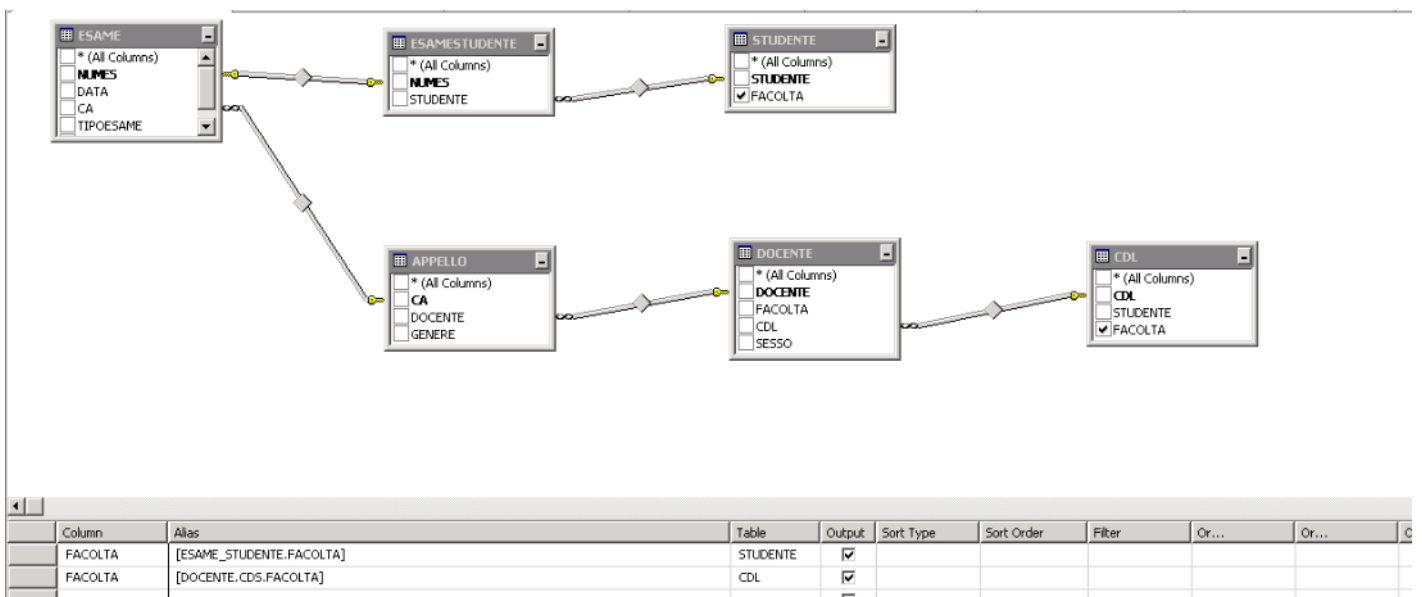
1) la facoltà dello studente (eventuale) che ha sostenuto l'esame **STUDENTE_ESAME.FACOLTA**



2) la facoltà del CDS del docente dell'appello dell'esame **DOCENTE.CDS.FACOLTA**:



IMPORTANTE: il join automaticamente aggiunto dal sistema tra STUDENTE e CDL deve ESSERE ELIMINATO (altrimenti andrei a consierare solo il caso degli esami degli studenti che sono rappresentati nel CDL del docente che ha fatto l'esame ...)



e procedo come in precedenza.

Concludiamo con il caso più difficile, ma che si presenta spesso in pratica ...

1) la facoltà dello **studente** (eventuale) che ha sostenuto l'esame **STUDENTE_ESAME.FACOLTA**

2) la facoltà dello **studente** rappresentante del CDS **DOCENTE.CDS.STUDENTE.FACOLTA**

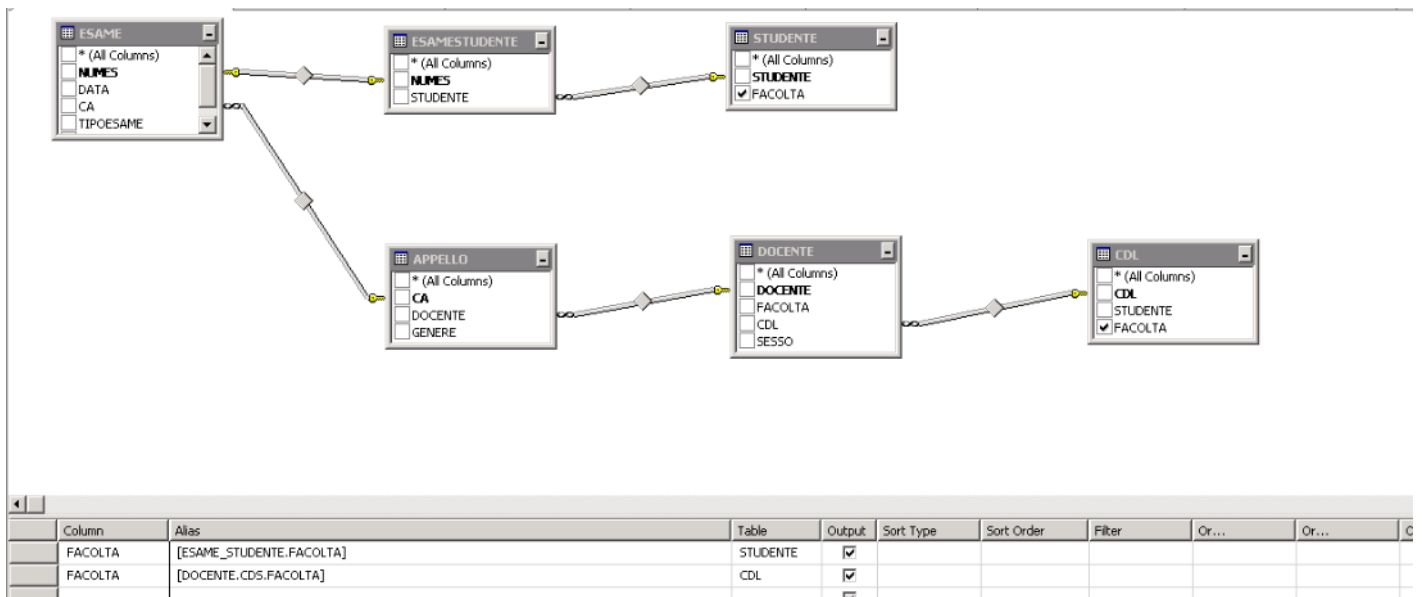
IMPORTANTE:

ho **due studenti**, che abbiamo detto **non necessariamente coincidenti**

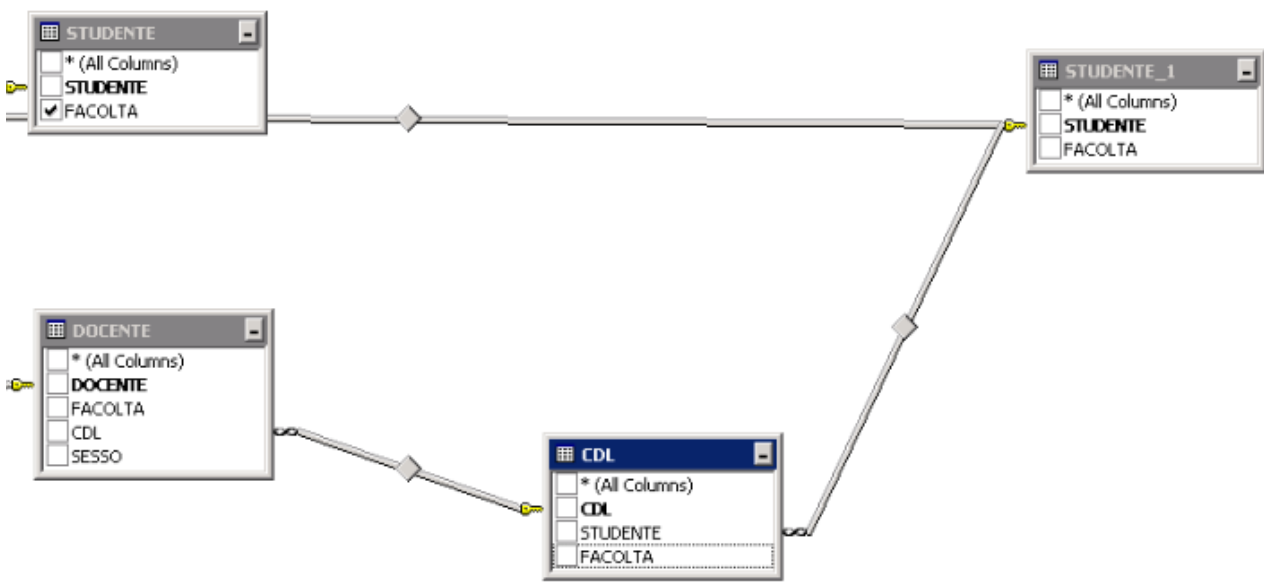
QUINDI

Devo mettere nella query **due volte** la tabella **STUDENTE**

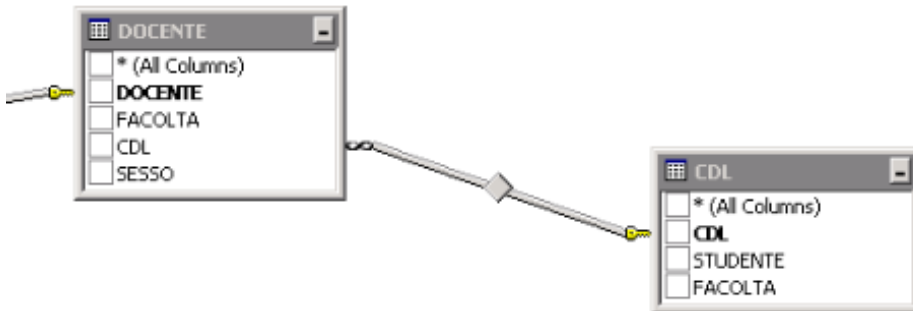
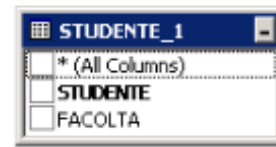
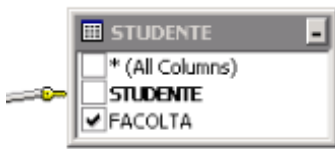
Si parte dalla precedente query, de-selezionando **DOCENTE.CDS.FACOLTA**



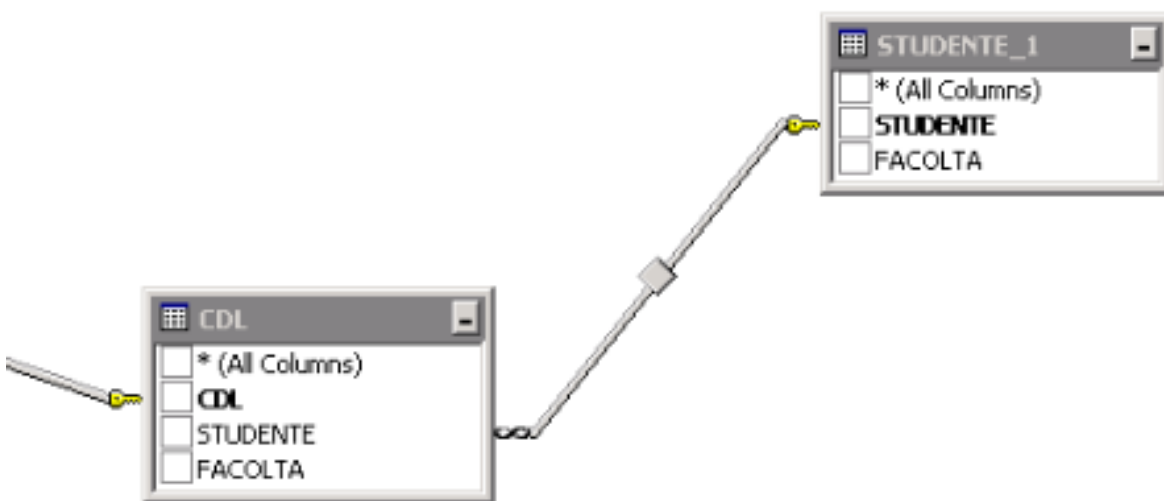
Si aggiunge un'altra tabella **STUDENTE** (il sistema gli assegna l'alias **STUDENTE_1**)



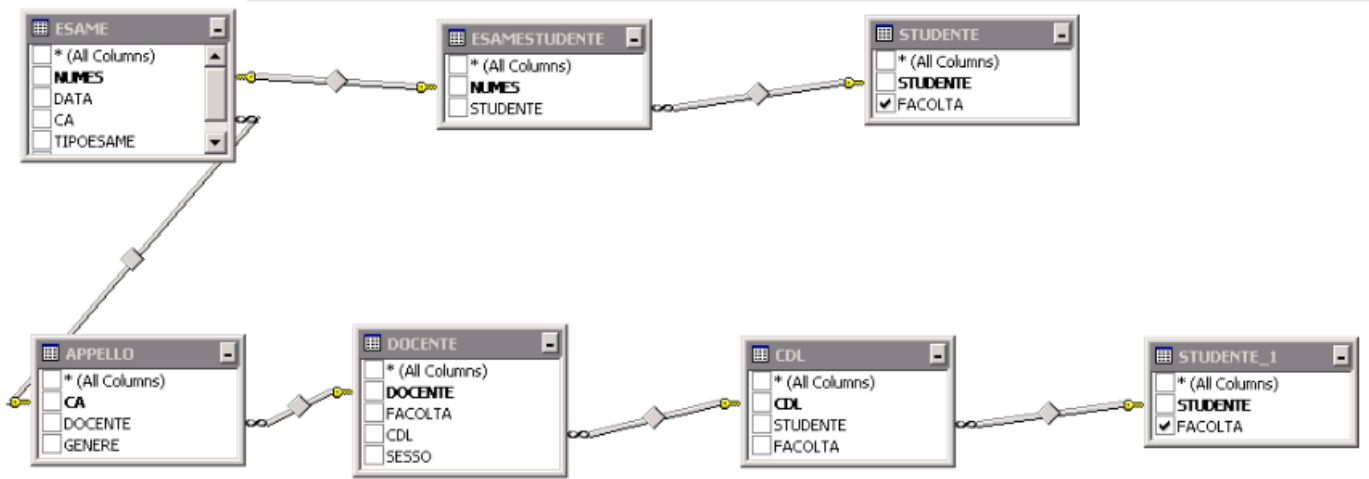
tolgo tutti i collegamenti automaticamente generati con **STUDENTE 1**



e definisco quello che mi interessa, cioè tra CDL e STUDENTE



seleziono quindi la FACOLTA di STUDENTE_1 che sarà quindi **DOCENTE.CDS.STUDENTE.FACOLTA**



| Column | Alias | Table | Output | Sort Type | Sort Order | Filter |
|---------|--------------------------------|------------|-------------------------------------|-----------|------------|--------|
| FACOLTA | [ESAME_STUDENTE.FACOLTA] | STUDENTE | <input checked="" type="checkbox"/> | | | |
| FACOLTA | [DOCENTE.CDS.STUDENTE.FACOLTA] | STUDENTE_1 | <input checked="" type="checkbox"/> | | | |

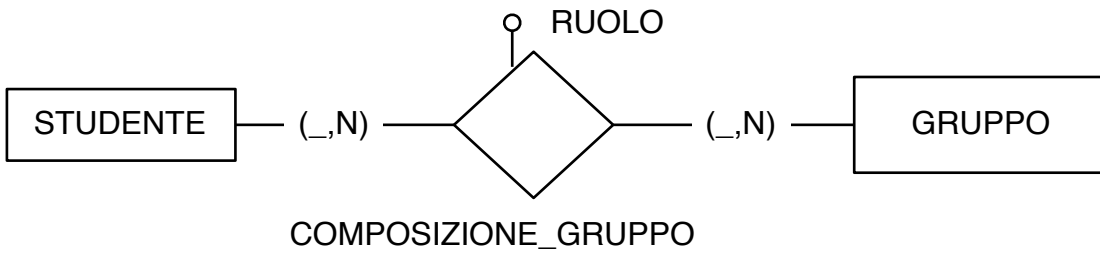
e procedo come in precedenza

Si considera anche COMPOSIZIONE_GRUPPO



COMPOSIZIONE_GRUPPO (GRUPPO:GRUPPO, STUDENTE:STUDENTE, RUOLO)

In E_R



NOTE:

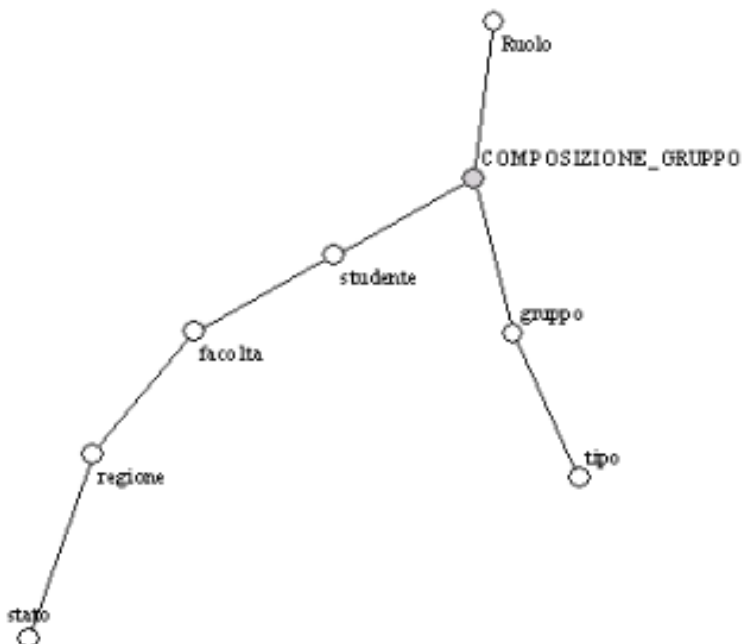
- 1) L'ER può essere disegnato *a pezzi*
- 2) Quando si ripete una entità già fatta in precedenza (STUDENTE e GRUPPO) non si deve ripetere tutto, basta il loro nome ...

FATTO: COMPOSIZIONE_GRUPPO

- 1) Si costruisce l'albero degli attributi come spiegato nelle dispense
http://www.dbgroup.unimo.it/SIA/SIA_2015_ProgettazioneConcettuale.pdf
(in particolare vedere attentamente l'esempio delle vendite che inizia a pag. 15)

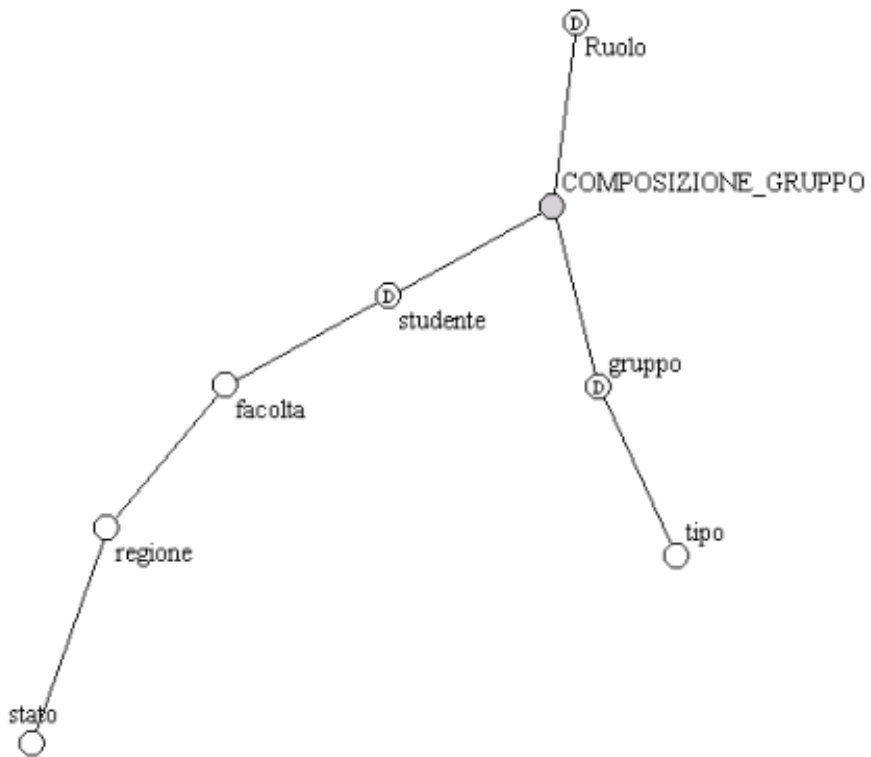


- 2) Editing dell'albero: SI AGGIUNGONO LE FD individuate (e non riportate nello schema E/R). Nel nostro esempio la FD in FACOLTA : REGIONE → STATO

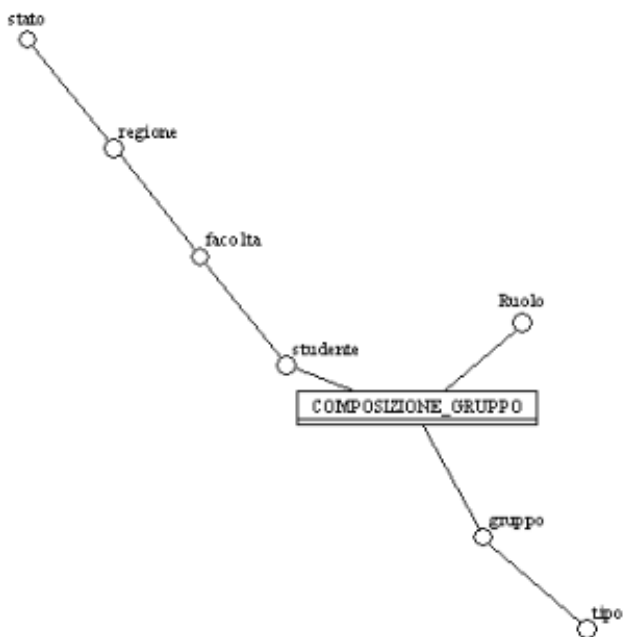


Si decide di non fare semplificazioni (innesti e potature)

DIMENSIONI: STUDENTE, GRUPPO e RUOLO

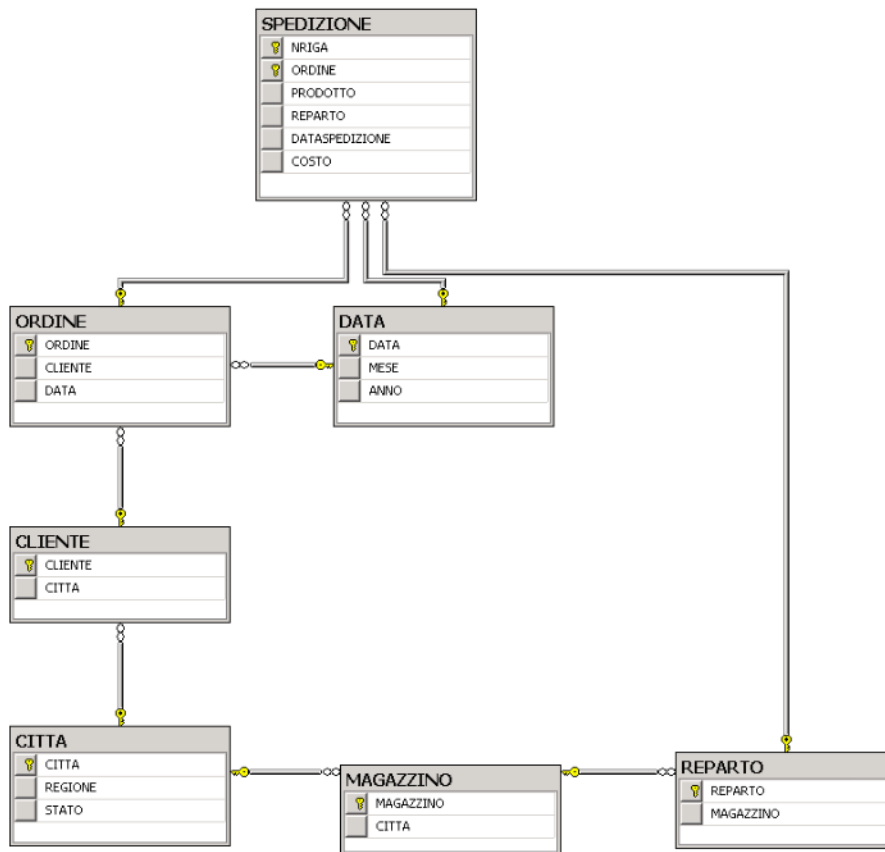


Si crea lo Schema di Fatto



Esercizio: Spedizione

Consideriamo un DBO con il seguente diagramma relazionale



A) DATA PROFILING

DF: REGIONE → STATO e **DF**: MESE → ANNO

B) SCHEMA RELAZIONALE

CITTA (CITTA, REGIONE, STATO)

DF: REGIONE → STATO

DATA (DATA, MESE, ANNO)

DF: MESE → ANNO

ORDINE (ORDINE, CLIENTE:CLIENTE, DATA:DATA)

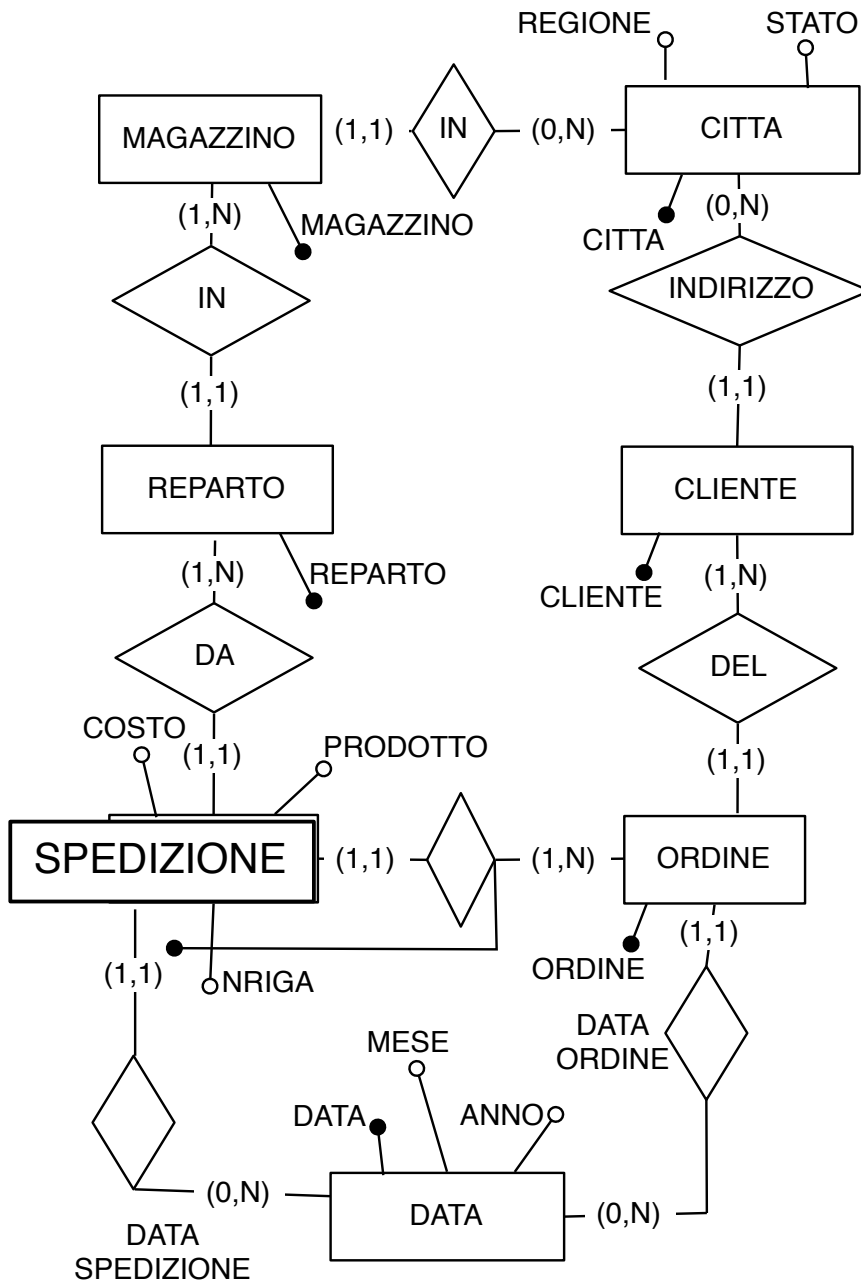
REPARTO (REPARTO, MAGAZZINO:MAGAZZINO)

MAGAZZINO (MAGAZZINO, CITTA:CITTA)

CLIENTE (CLIENTE, CITTA:CITTA)

SPEDIZIONE (NRIGA, ORDINE:ORDINE,
PRODOTTO,
REPARTO:REPARTO
DATASPEDIZIONE:DATA,
COSTO)

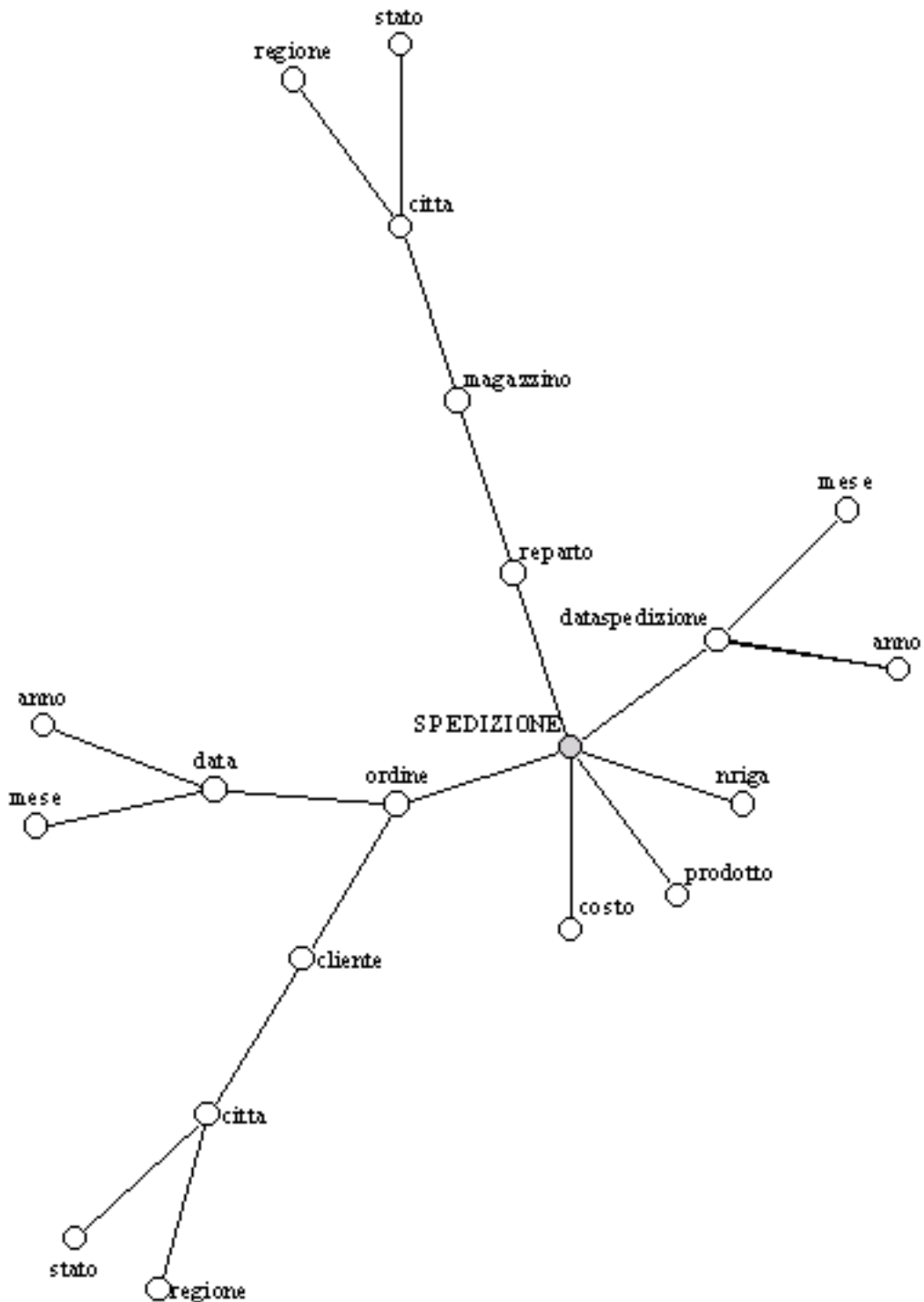
C) SCHEMA ER



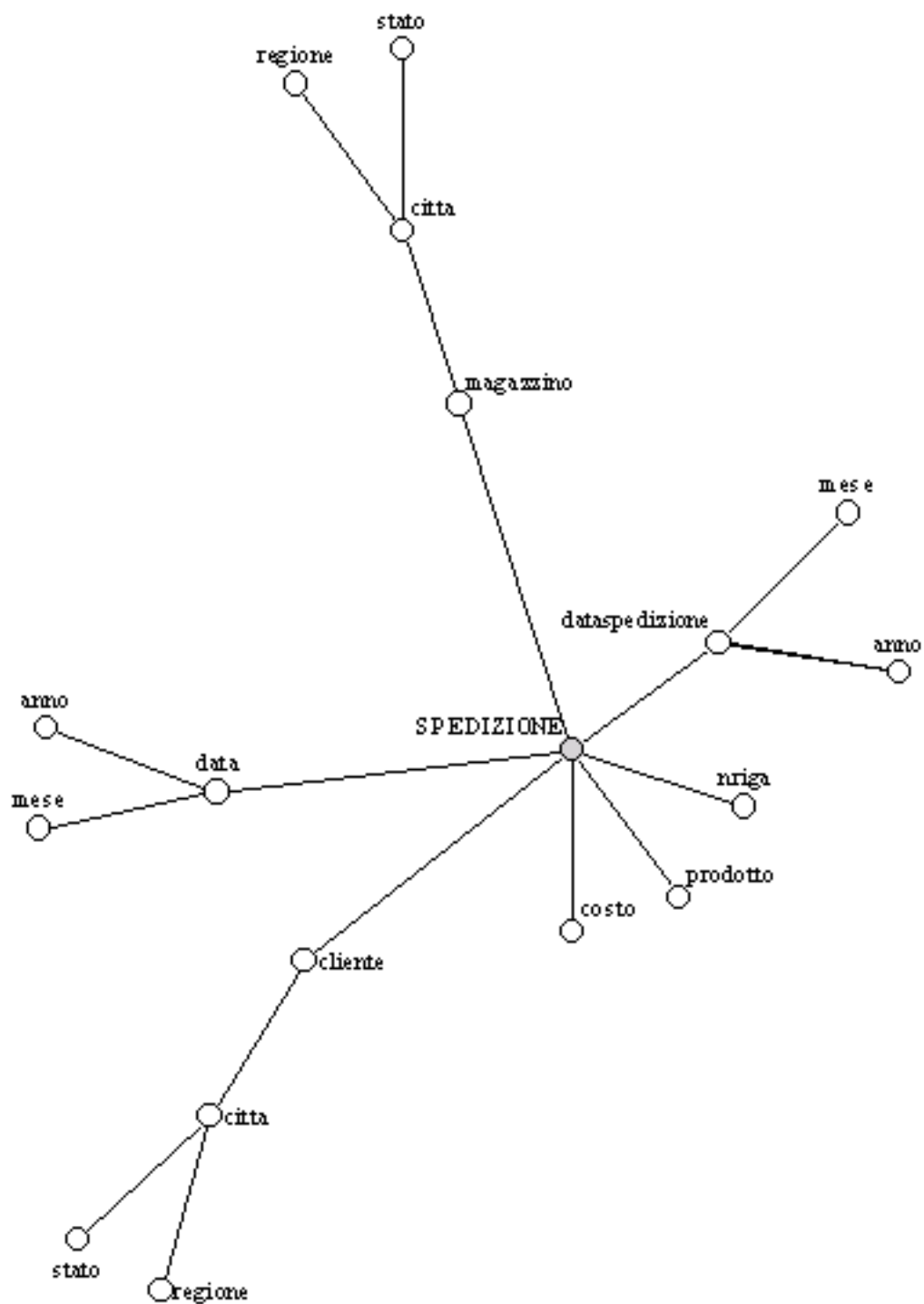
D) PROGETTO CONCETTUALE

fatto SPEDIZIONE con **dimensioni**
{PRODOTTO,MAGAZZINO,CLIENTE,DATASPED}

Albero degli attributi iniziale



INNESTO su REPARTO ed ORDINE



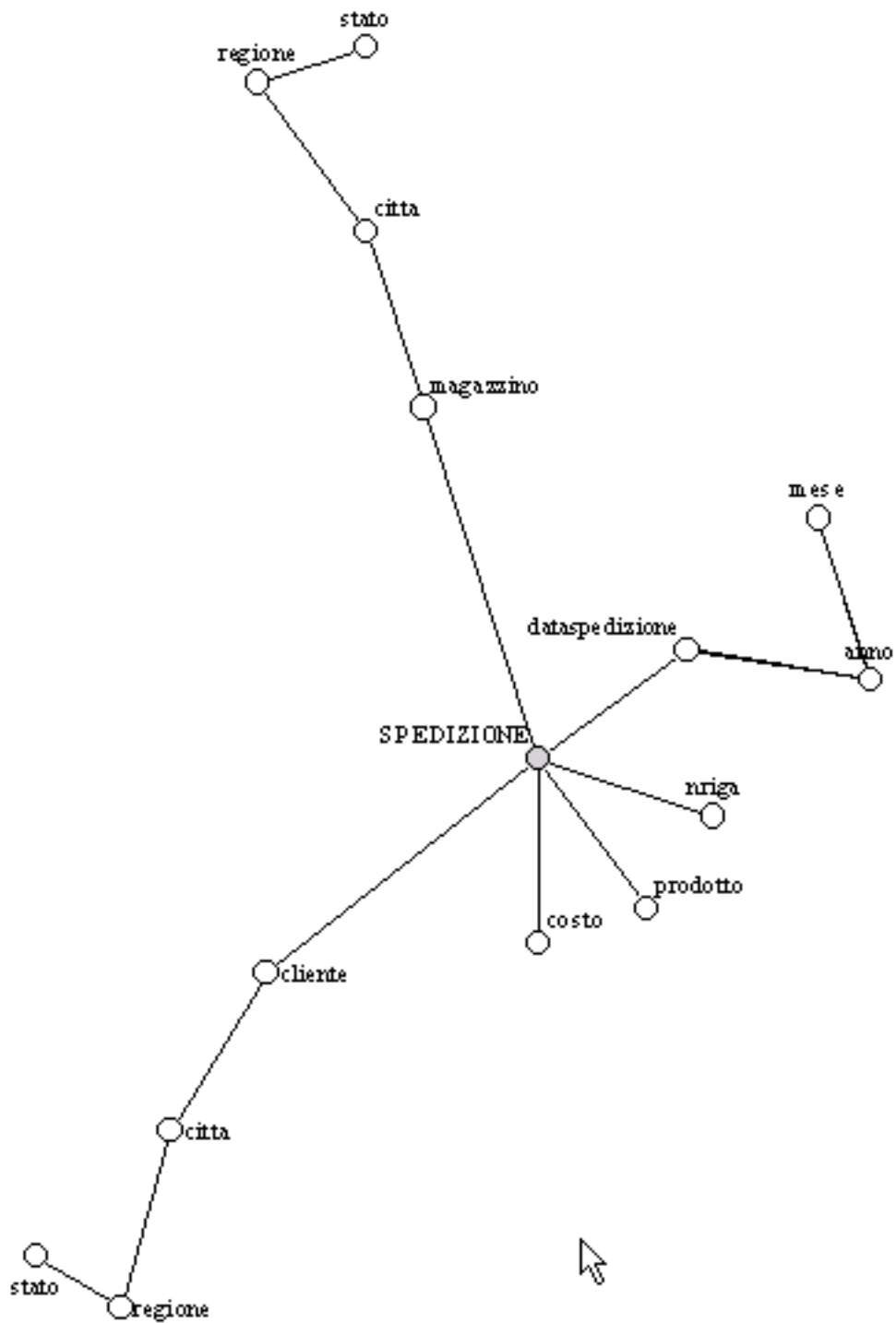
Anche DATA dell'ORDINE viene POTATA

Si aggiungono le seguenti dipendenze funzionali, non presenti nello schema ER:

REGIONE → STATO

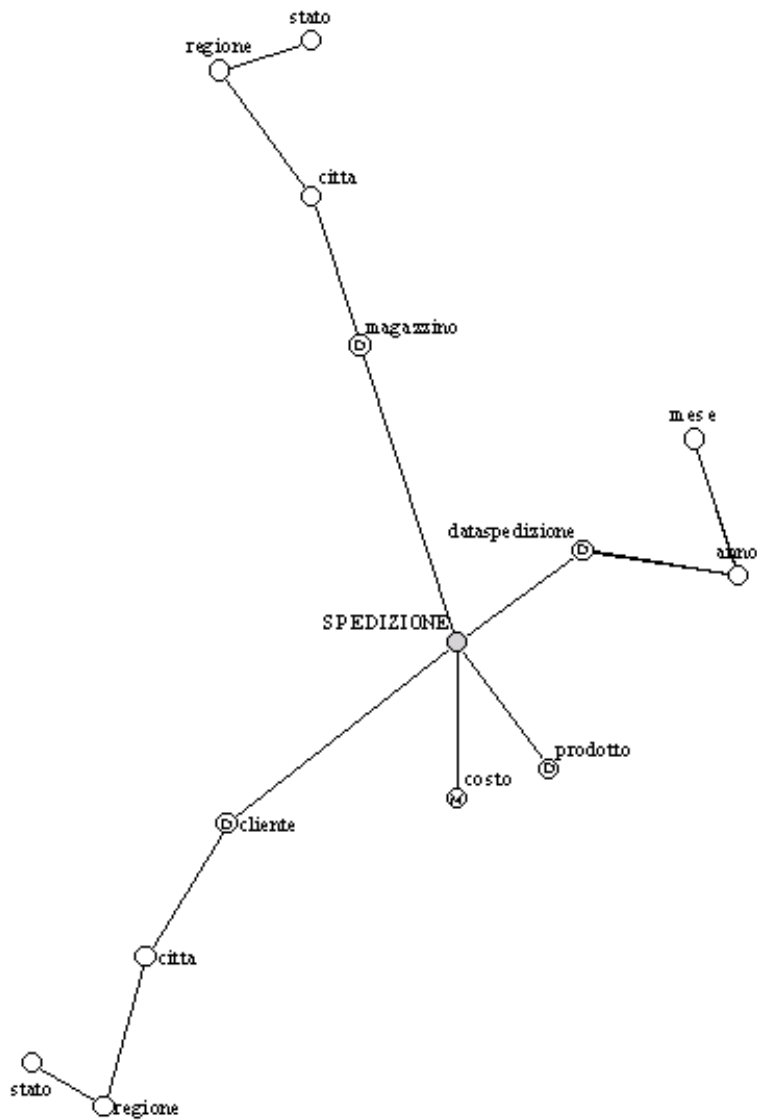
MESE → ANNO

Si ottiene



Viene POTATA anche NRIGA

Scelta delle dimensioni



Viene quindi creato lo schema di fatto, considerando una condivisione su CITTA (le misure e le relative aggregabilità verranno analizzate ed indicate sullo schema di fatto **nella seconda consegna**).

