

Basi di dati

Capitolo 8: Progettazione logica

1/75

Obiettivo della progettazione logica

- "tradurre" lo schema concettuale in uno schema logico che rappresenti gli stessi dati in maniera corretta ed efficiente

3/75

Non si tratta di una pura e semplice traduzione

- alcuni aspetti non sono direttamente rappresentabili
- è necessario considerare le prestazioni

5/75

Requisiti della base di dati

Progettazione
concettuale

Schema concettuale

Progettazione
logica

Schema logico

Progettazione
fisica

Schema fisico

2/75

Dati di ingresso e uscita

- **Ingresso:**
 - schema concettuale
 - informazioni sul carico applicativo
 - modello logico
- **Uscita:**
 - schema logico
 - documentazione associata

4/75

Carico
applicativo

Schema concettuale
E-R

Ristrutturazione dello
schema E-R

Modello
logico

Schema E-R
ristrutturato

Traduzione nel
modello logico

Schema
logico

6/75

Ristrutturazione schema E-R

- Motivazioni:
 - semplificare la traduzione
 - "ottimizzare" le prestazioni
- Osservazione:
 - uno schema E-R ristrutturato non è (più) uno schema concettuale nel senso stretto del termine

7/75

Prestazioni, approssimate

- Consideriamo:
 - "indicatori" dei parametri che regolano le prestazioni
- spazio:
 - numero di occorrenze previste
- tempo:
 - numero di occorrenze (di entità e relationship) visitate durante un'operazione

9/75

Tavola dei volumi

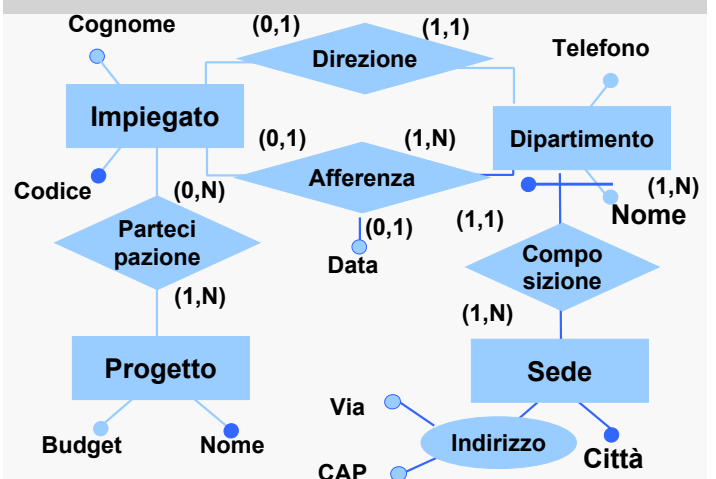
Concetto	Tipo	Volume
Sede	E	10
Dipartimento	E	80
Impiegato	E	2000
Progetto	E	500
Composizione	R	80
Afferenza	R	1900
Direzione	R	80
Partecipazione	R	6000

11/75

Prestazioni?

- Per ottimizzare il risultato abbiamo bisogno di analizzare le prestazioni a questo livello
- Ma:
 - le prestazioni non sono valutabili con precisione su uno schema concettuale!

8/75

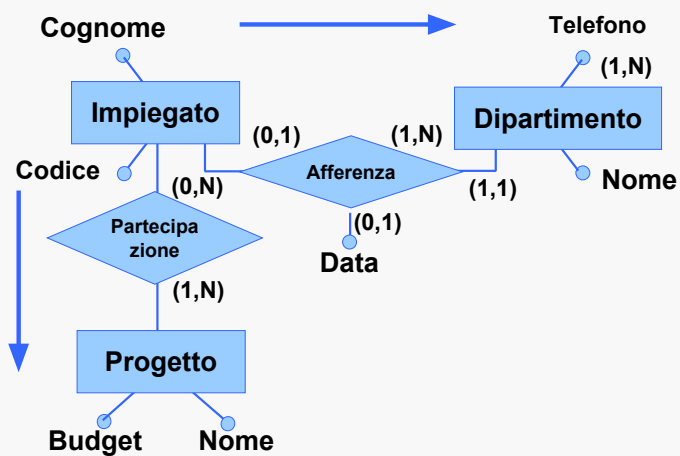


10/75

Esempio di valutazione di costo

- Operazione:
 - trova tutti i dati di un impiegato, del dipartimento nel quale lavora e dei progetti ai quali partecipa
- Si costruisce una tavola degli accessi basata su uno schema di navigazione

12/75



13/75

Tavola degli accessi

Concetto	Costrutto	Accessi	Tipo
Impiegato	Entità	1	L
Afferenza	Relationship	1	L
Dipartimento	Entità	1	L
Partecipazione	Relationship	3	L
Progetto	Entità	3	L

14/75

Attività della ristrutturazione

- Analisi delle ridondanze
- Eliminazione delle generalizzazioni
- Partizionamento/accorpamento di entità e relationship
- Scelta degli identificatori primari

15/75

Analisi delle ridondanze

- Una ridondanza in uno schema E-R è una informazione significativa ma derivabile da altre
- in questa fase si decide se eliminare le ridondanze eventualmente presenti o mantenerle (o anche di introdurne di nuove)

16/75

Ridondanze

- Vantaggi
 - semplificazione delle interrogazioni
- Svantaggi
 - appesantimento degli aggiornamenti
 - maggiore occupazione di spazio

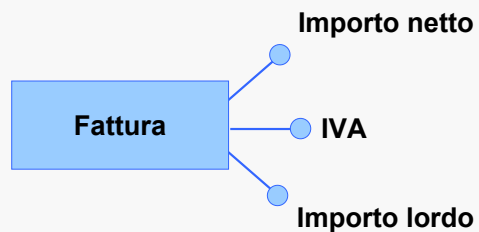
17/75

Forme di ridondanza in uno schema E-R

- attributi derivabili:
 - da altri attributi della stessa entità (o relationship)
 - da attributi di altre entità (o relationship)
- relationship derivabili dalla composizione di altre (più in generale: cicli di relationship)

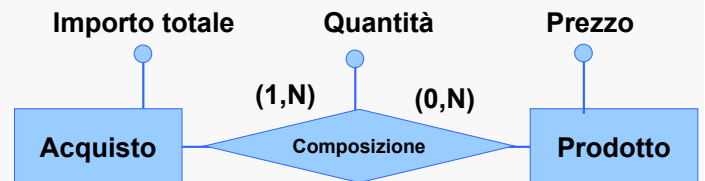
18/75

Attributo derivabile



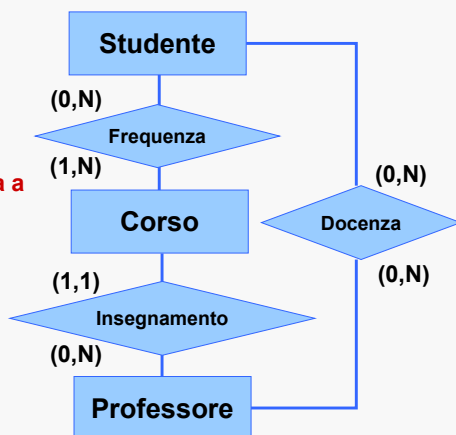
19/75

Attributo derivabile da altra entità



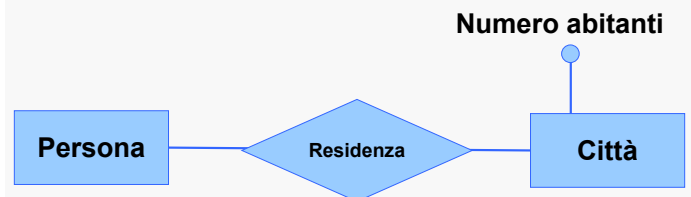
20/75

Ridondanza dovuta a ciclo



21/75

Analisi di una ridondanza



22/75

Concetto	Tipo	Volume
Città	E	200
Persona	E	1000000
Residenza	R	1000000

- **Operazione 1:** memorizza una nuova persona con la relativa città di residenza (500 volte al giorno)
- **Operazione 2:** stampa tutti i dati di una città (incluso il numero di abitanti) (2 volte al giorno)

23/75

Presenza di ridondanza

Operazione 1

Concetto	Costrutto	Accessi	Tipo
Persona	Entità	1	S
Residenza	Relazione	1	S
Città	Entità	1	L
Città	Entità	1	S

Operazione 2

Concetto	Costrutto	Accessi	Tipo
Città	Entità	1	L

24/75

Assenza di ridondanza

Operazione 1

Concetto	Costrutto	Accessi	Tipo
Persona	Entità	1	S
Residenza	Relazione	1	S

Operazione 2

Concetto	Costrutto	Accessi	Tipo
Città	Entità	1	L
Residenza	Relazione	5000	L

25/75

Presenza di ridondanza

- Costi:
 - Operazione 1: 1500 accessi in scrittura e 500 accessi in lettura al giorno
 - Operazione 2: trascurabile.
- Contiamo doppi gli accessi in scrittura
 - Totale di 3500 accessi al giorno

26/75

Assenza di ridondanza

- Costi:
 - Operazione 1: 1000 accessi in scrittura
 - Operazione 2: 10000 accessi in lettura al giorno
- Contiamo doppi gli accessi in scrittura
 - Totale di 12000 accessi al giorno

27/75

Attività della ristrutturazione

- Analisi delle ridondanze
- Eliminazione delle generalizzazioni
- Partizionamento/accorpamento di entità e relationship
- Scelta degli identificatori primari

28/75

Eliminazione delle gerarchie

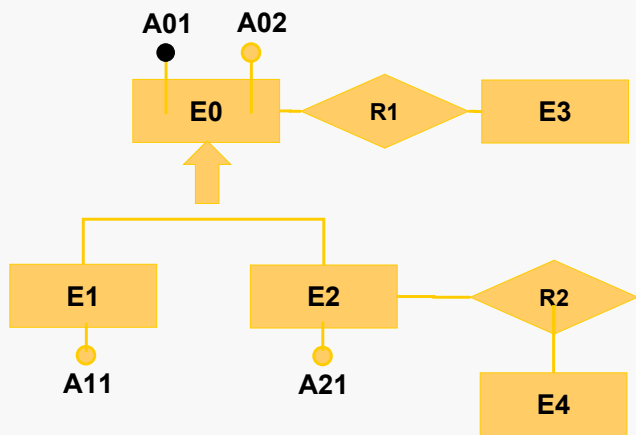
- il modello relazionale non può rappresentare direttamente le generalizzazioni
- entità e relationship sono invece direttamente rappresentabili
 - si eliminano perciò le gerarchie, sostituendole con entità e relationship

29/75

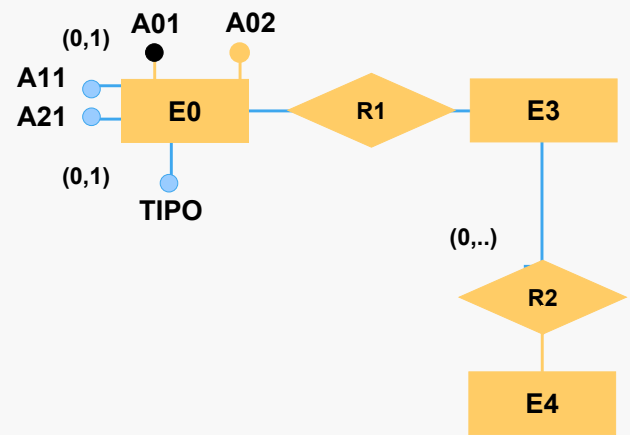
Tre possibilità

1. accorpamento delle figlie della generalizzazione nel genitore
2. accorpamento del genitore della generalizzazione nelle figlie
3. sostituzione della generalizzazione con relationship

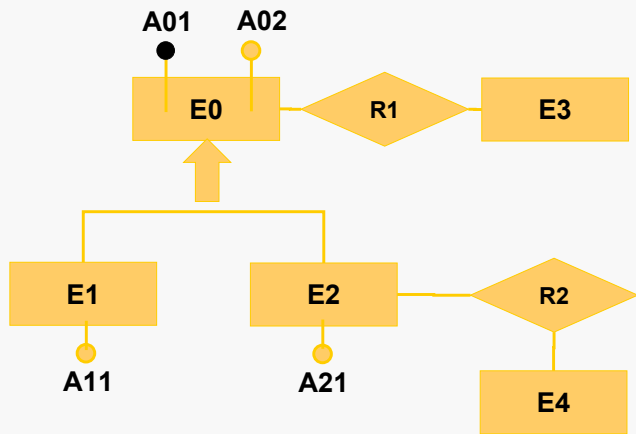
30/75



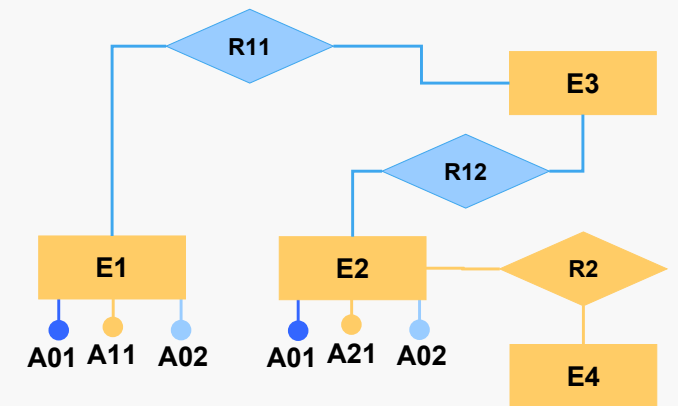
31/75



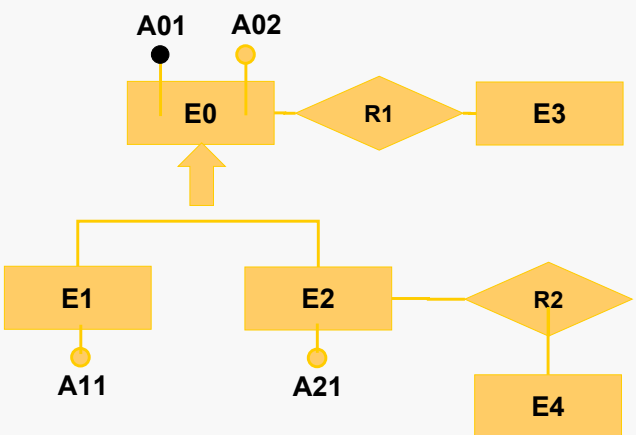
32/75



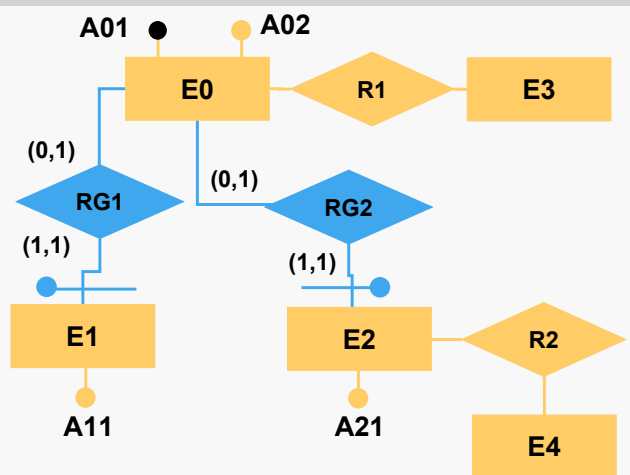
33/75



34/75



35/75



36/75

- la scelta fra le alternative si può fare con metodo simile a quello visto per l'analisi delle ridondanze (però non basato solo sul numero degli accessi)

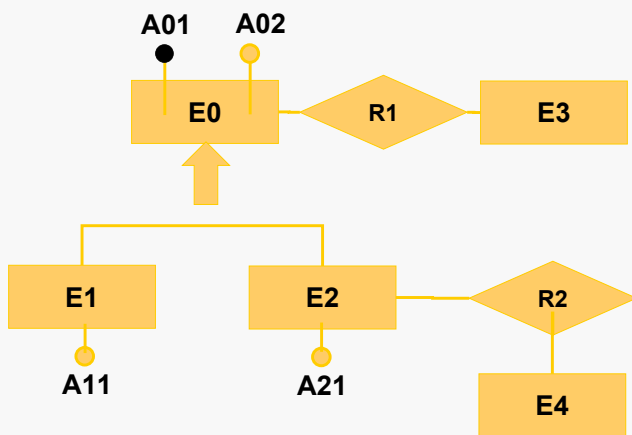
- è possibile seguire alcune semplici regole generali

37/75

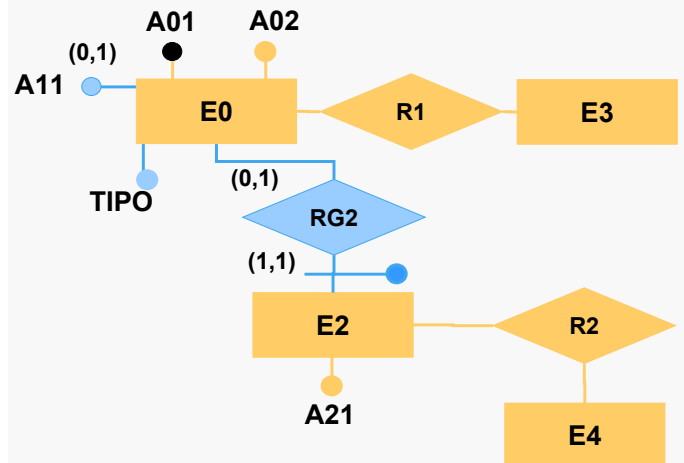
- conviene se gli accessi al padre e alle figlie sono contestuali
- conviene se gli accessi alle figlie sono distinti
- conviene se gli accessi alle entità figlie sono separati dagli accessi al padre

- sono anche possibili soluzioni "ibride", soprattutto in gerarchie a più livelli

38/75



39/75



40/75

Attività della ristrutturazione

- Analisi delle ridondanze
- Eliminazione delle generalizzazioni
- Partizionamento/accorpamento di entità e relationship
- Scelta degli identificatori primari

41/75

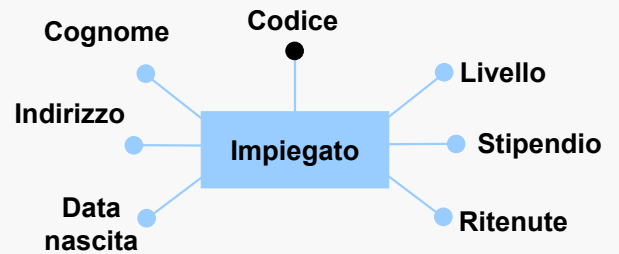
- Ristrutturazioni effettuate per rendere più efficienti le operazioni in base a un semplice principio
- Gli accessi si riducono:
 - separando attributi di un concetto che vengono acceduti separatamente
 - raggruppando attributi di concetti diversi acceduti insieme

42/75

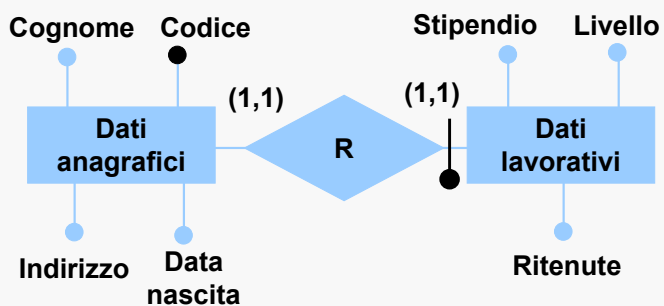
Ristrutturazioni, casi principali

- partizionamento verticale di entità
- partizionamento orizzontale di relationship
- eliminazione di attributi multivalore
- accorpamento di entità/ relationship

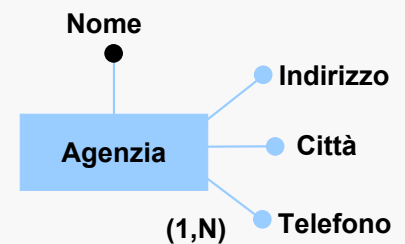
43/75



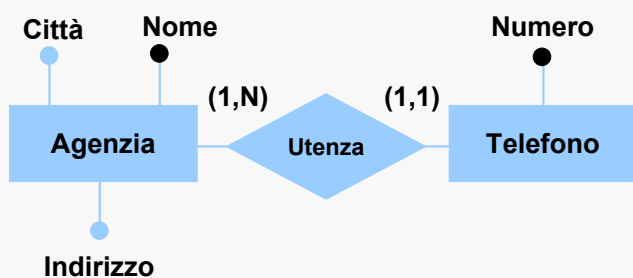
44/75



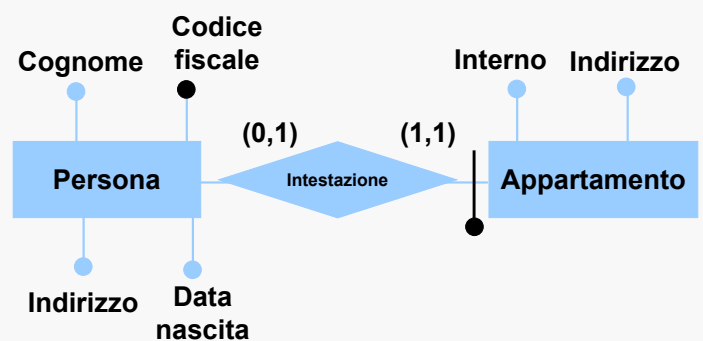
45/75



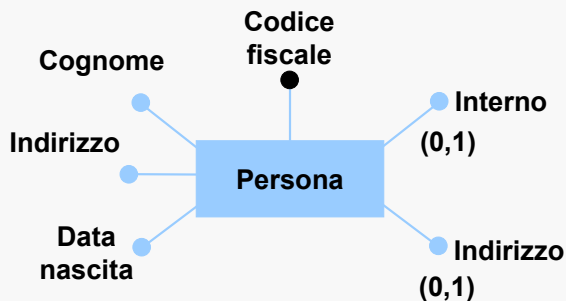
46/75



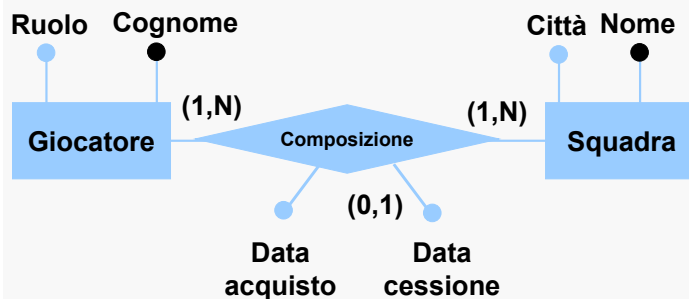
47/75



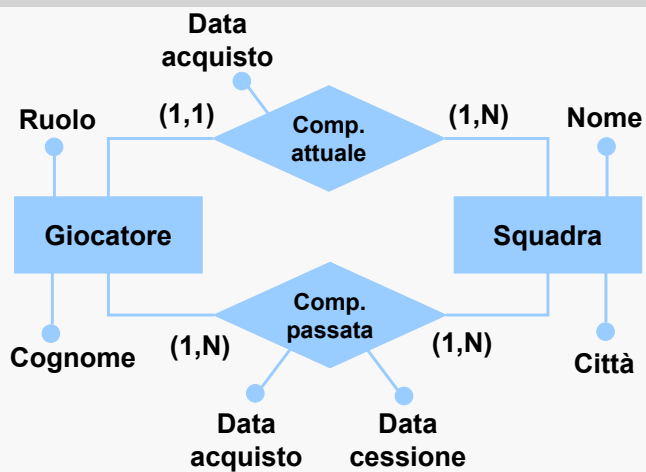
48/75



49/75



50/75



51/75

Attività della ristrutturazione

- Analisi delle ridondanze
- Eliminazione delle generalizzazioni
- Partizionamento/accorpamento di entità e relationship
- Scelta degli identificatori principali

52/75

Scelta degli identificatori principali

- operazione indispensabile per la traduzione nel modello relazionale
- Criteri
 - assenza di opzionalità
 - semplicità
 - utilizzo nelle operazioni più frequenti o importanti

53/75

Se nessuno degli identificatori soddisfa i requisiti visti?

Si introducono nuovi attributi (**codici**) contenenti valori speciali generati appositamente per questo scopo

54/75

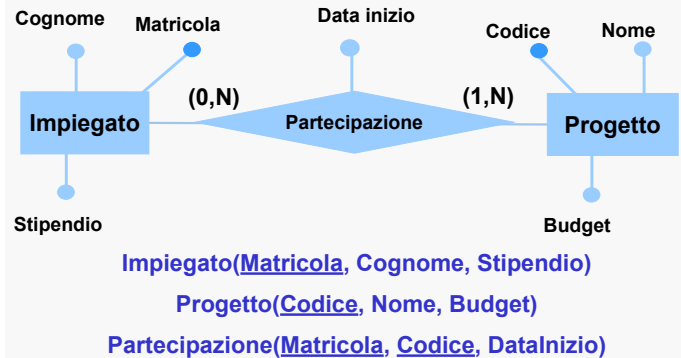
Traduzione verso il modello relazionale

• idea di base:

- le entità diventano relazioni sugli stessi attributi
- le relationship diventano relazioni sugli identificatori delle entità coinvolte (più gli attributi propri)

55/75

Entità e relationship molti a molti



56/75

Entità e relationship molti a molti

Impiegato(Matricola, Cognome, Stipendio)
 Progetto(Codice, Nome, Budget)
 Partecipazione(Matricola, Codice, DataInizio)

- con vincoli di integrità referenziale fra
 - Matricola in **Partecipazione** e (la chiave di) **Impiegato**
 - Codice in **Partecipazione** e (la chiave di) **Progetto**

57/75

Nomi più espressivi per gli attributi della chiave della relazione che rappresenta la relationship

Impiegato(Matricola, Cognome, Stipendio)
 Progetto(Codice, Nome, Budget)
 Partecipazione(Matricola, Codice, DataInizio)
 Partecipazione(Impiegato, Progetto, DataInizio)

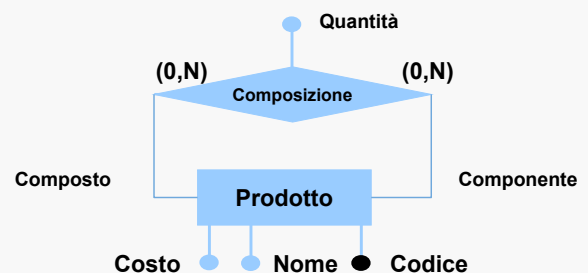
58/75

Nota

- La traduzione non riesce a tener conto delle cardinalità minime delle relationship molti a molti (se non con vincoli di CHECK complessi e poco usati)

59/75

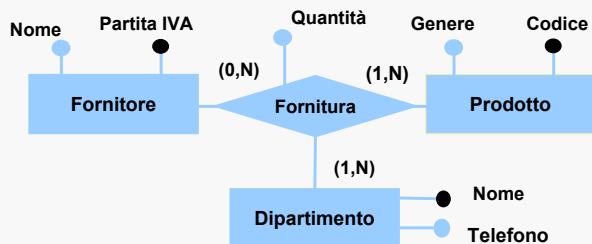
Relationship ricorsive



Prodotto(Codice, Nome, Costo)
 Composizione(Composto, Componente, Quantità)

60/75

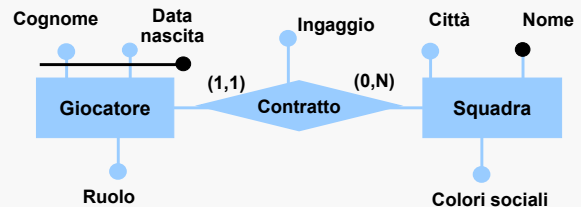
Relationship n-arie



Fornitore(PartitaIVA, Nome)
 Prodotto(Codice, Genere)
 Dipartimento(Nome, Telefono)
 Fornitura(Fornitore, Prodotto, Dipartimento, Quantità)

61/75

Relationship uno a molti



Giocatore(Cognome, DataNascita, Ruolo)
 Contratto(CognGiocatore, DataNascG, Squadra, Ingaggio)
 Squadra(Nome, Città, ColoriSociali)

- corretto?

62/75

Soluzione più compatta

Giocatore(Cognome, DataNascita, Ruolo)
 Contratto(CognGiocatore, DataNascG, Squadra, Ingaggio)
 Squadra(Nome, Città, ColoriSociali)

Giocatore(Cognome, DataNasc, Ruolo, Squadra, Ingaggio)
 Squadra(Nome, Città, ColoriSociali)

- con vincolo di integrità referenziale fra Squadra in Giocatore e la chiave di Squadra
- se la cardinalità minima della relationship è 0, allora Squadra in Giocatore deve ammettere valore nullo

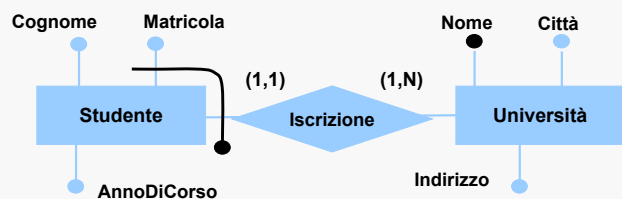
63/75

Nota

- La traduzione riesce a rappresentare efficacemente la cardinalità minima della partecipazione che ha 1 come cardinalità massima:
 - 0 : valore nullo ammesso
 - 1 : valore nullo non ammesso

64/75

Entità con identificazione esterna

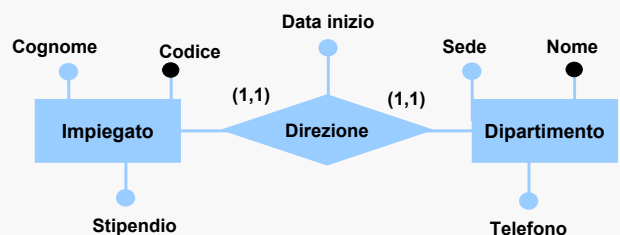


Studente(Matricola, Università, Cognome, AnnoDiCorso)
 Università(Nome, Città, Indirizzo)

- con vincolo ...

65/75

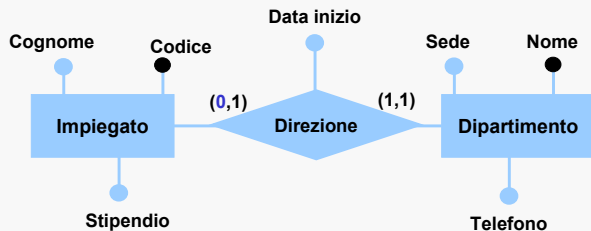
Relationship uno a uno



- varie possibilità:
 - fondere da una parte o dall'altra
 - fondere tutto?

66/75

Una possibilità privilegiata



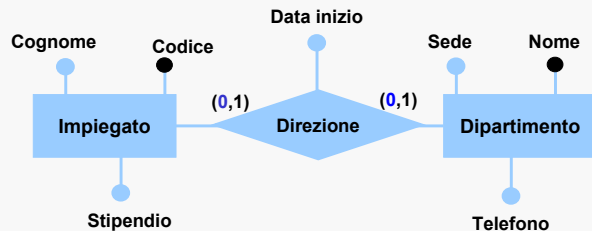
Impiegato (Codice, Cognome, Stipendio)

Dipartimento (Nome, Sede, Telefono, Direttore, InizioD)

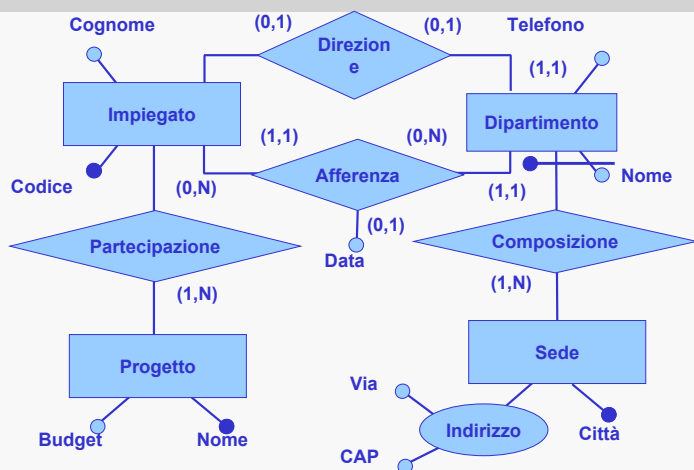
- con vincolo di integrità referenziale, senza valori nulli

67/75

Un altro caso



68/75



69/75

Schema finale

Impiegato(Codice, Cognome, Dipartimento, Sede, Data*)

Dipartimento(Nome, Città, Telefono, Direttore*)

Sede(Città, Via, CAP)

Progetto(Nome, Budget)

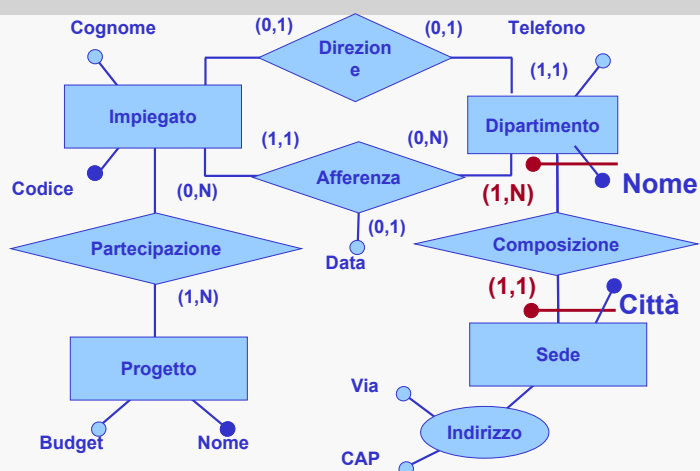
Partecipazione(Impiegato, Progetto)

70/75

Attenzione

- Differenze apparentemente piccole in cardinalità e indicatori possono cambiare di molto il significato ...

71/75

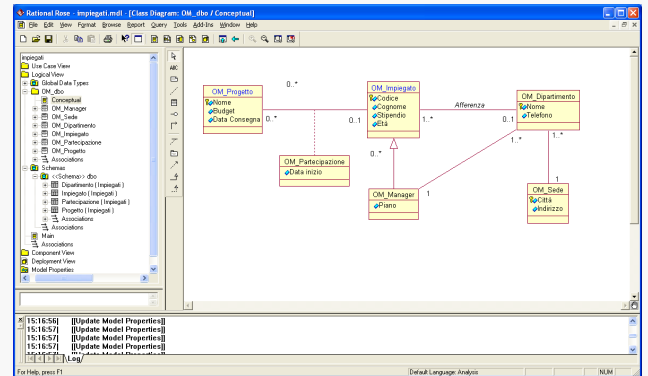


72/75

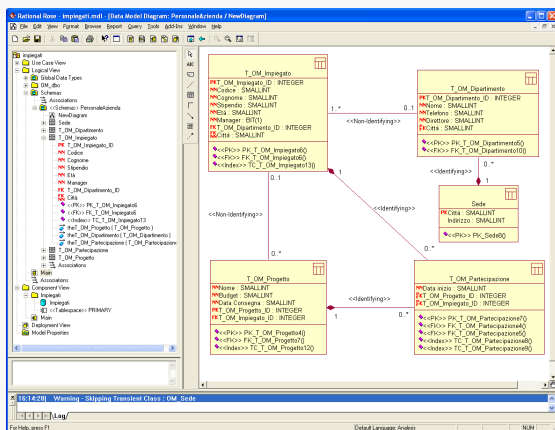
Strumenti di supporto

- Esistono sul mercato prodotti CASE che forniscono un supporto a tutte le fasi della progettazione di basi di dati

73/75



74/75



75/75