



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Progettazione Concettuale

Basi di Dati

Corso di Laurea in Informatica per il Management

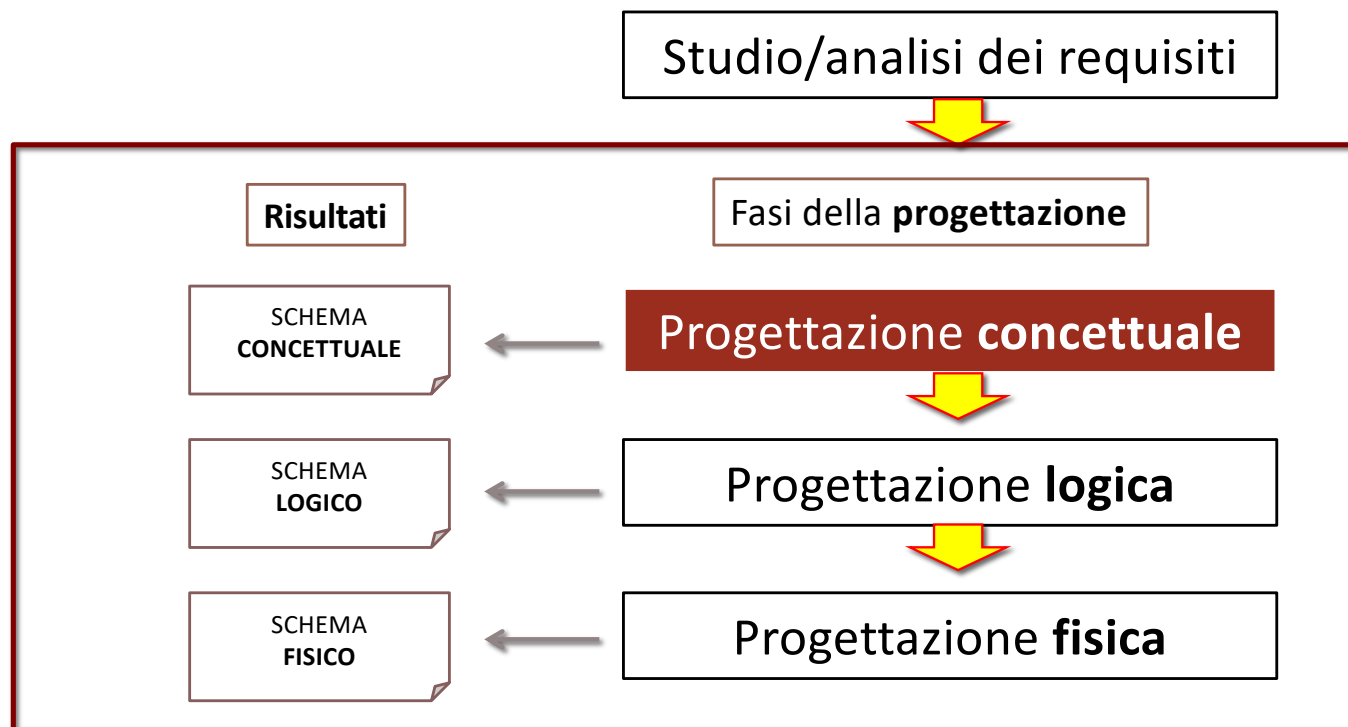
Alma Mater Studiorum - Università di Bologna

Prof. Marco Di Felice

Dipartimento di Informatica – Scienza e Ingegneria

marco.difelice3@unibo.it

Progettazione concettuale



Progettazione concettuale

La costruzione di uno schema concettuale deve tenere conto di alcune **proprietà generali** che ne determinano la **qualità**:

- **Correttezza** → **utilizzo corretto** (sintattico/semantico) **dei costrutti** del modello E-R.
- **Completezza** → **rappresentazione di tutti i dati di interesse** (e delle operazioni) descritti nel documento di specifica.

Progettazione concettuale

Come garantire tali proprietà?

In generale dipendono dal progettista, ma esistono alcune **metodologie di progettazione concettuale ...**

- **Strategie di progettazione**
- Pattern di progettazione
- Analisi di prestazione

Progettazione concettuale

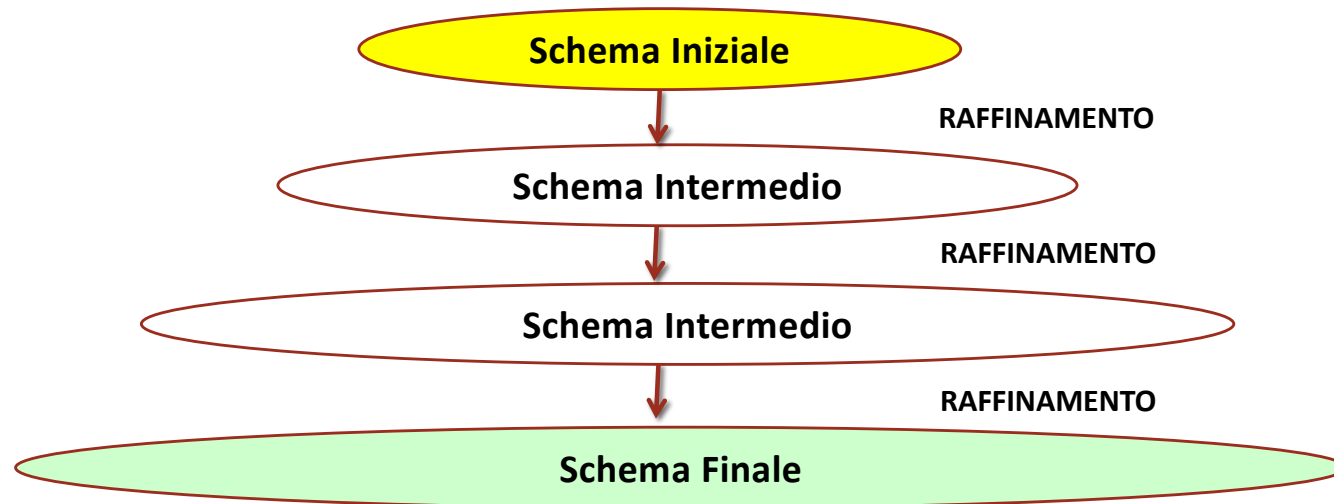


In generale, il documento di specifica potrebbe essere molto complesso e denso di contenuti ... **da dove partire per la costruzione del modello E-R?**

- **Strategia top-down**
- Strategia bottom-up
- Strategia inside-out
- Strategia mista

Progettazione concettuale

Nella progettazione **top-down**, lo schema concettuale viene ottenuto attraverso una **serie di raffinamenti successivi a partire da uno schema iniziale** molto astratto.



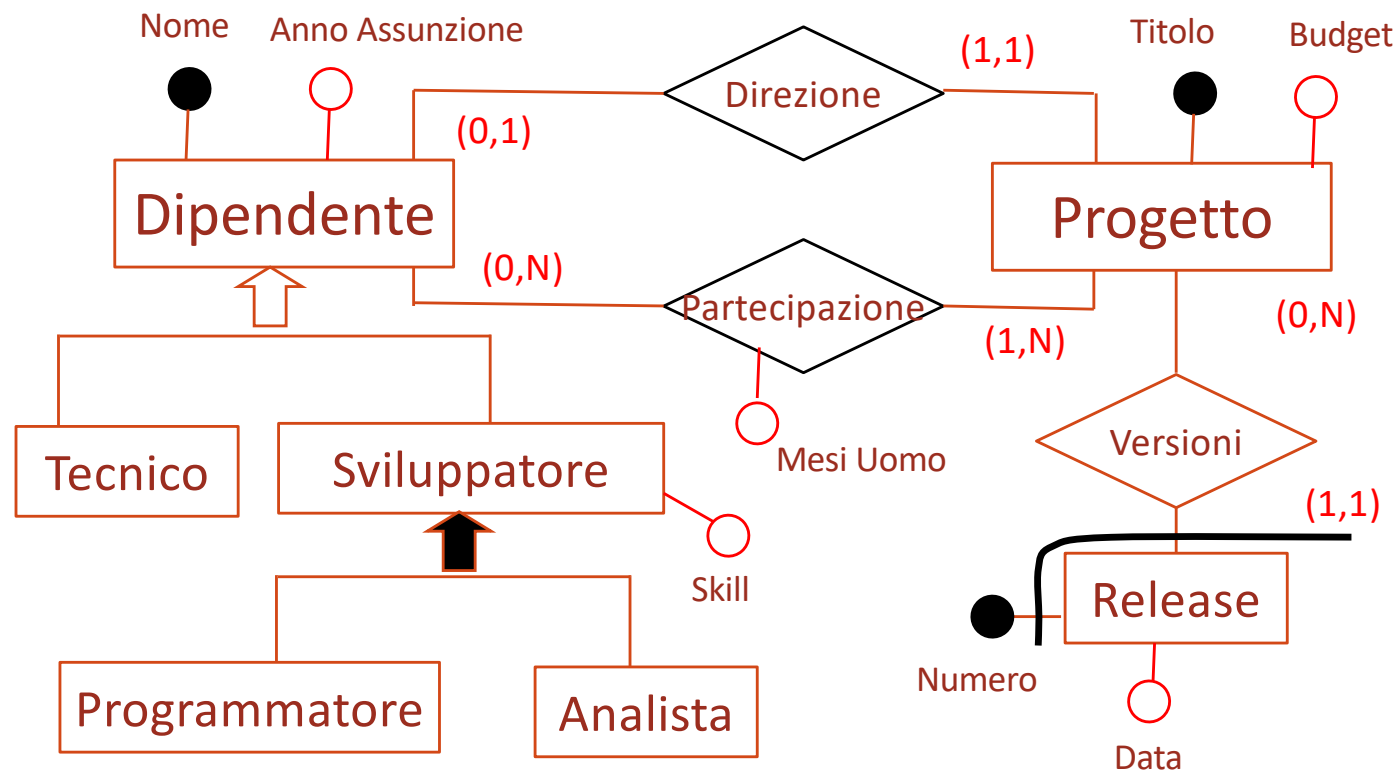
Progettazione concettuale

ANALISI REQUISITI



Si vuole realizzare una base di dati di una Software House, di cui si vogliono rappresentare i dati dei dipendenti e dei progetti. Ogni dipendente e' identificato da un nome, ed ha una stipendio e data di assunzione. I dipendenti possono far parte di progetti software. E' consentito ad un dipendente di partecipare a piu' di un progetto, allocando mesi-uomo su ciascun progetto. Inoltre, ogni progetto ha un direttore unico tra i dipendenti che vi partecipano. Ogni progetto ha un titolo univoco, un budget, e puo' disporre di diverse release. Ad ogni release di un progetto e' associato un numero progressivo ed una data. Tra i dipendenti, si vogliono gestire i dati dei tecnici e degli sviluppatori. Di ogni sviluppatore, si vuole tenere traccia delle skill specifiche. Gli sviluppatori sono ulteriormente suddivisi in Programmatori ed Analisti.

Progettazione concettuale



Progettazione concettuale

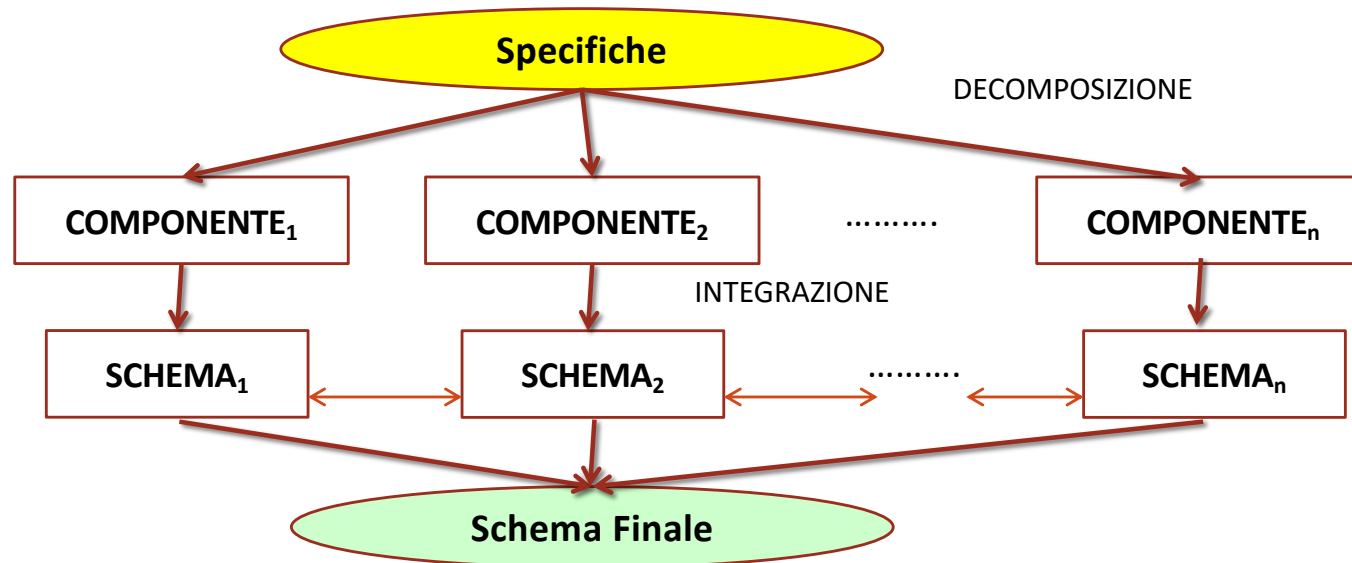


In generale, il documento di specifica potrebbe essere molto complesso e denso di contenuti ... **da dove partire per la costruzione del modello E-R?**

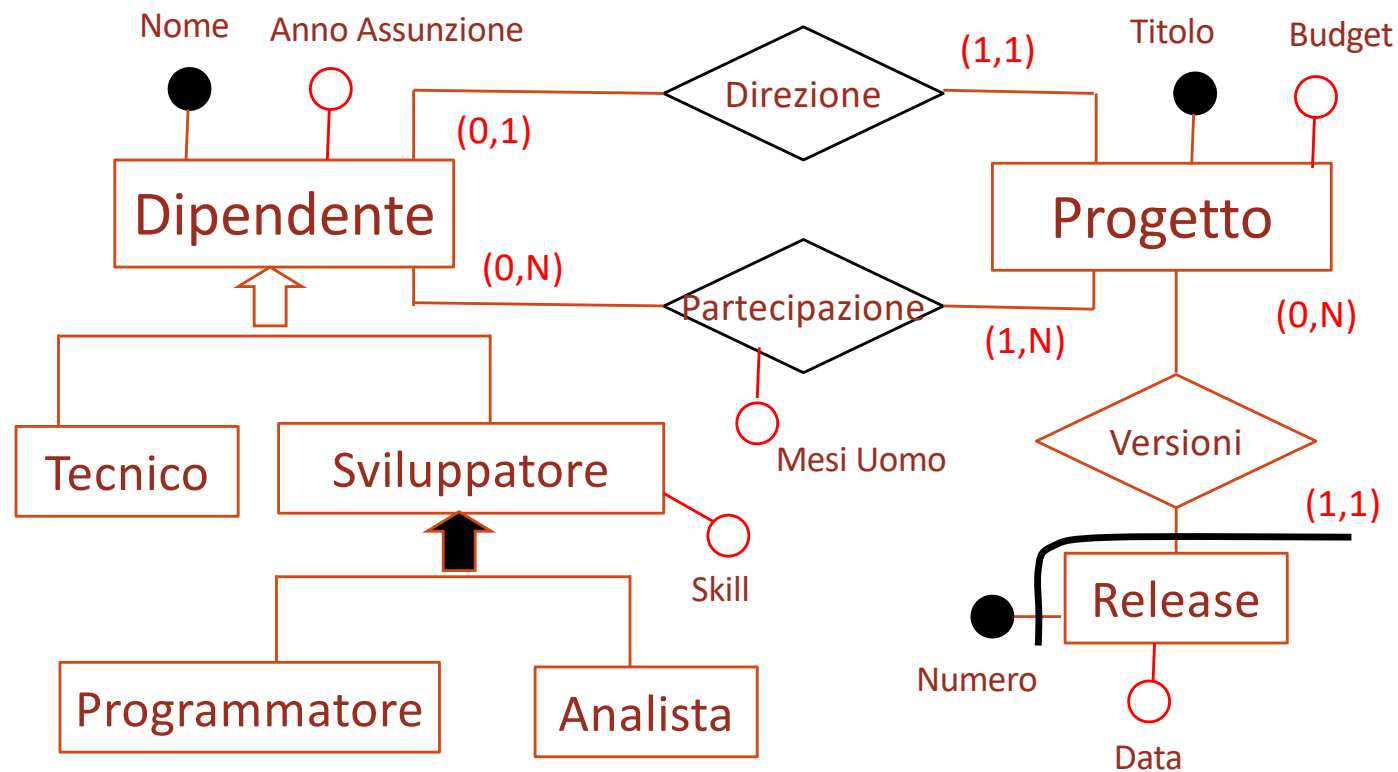
- Strategia top-down
- **Strategia bottom-up**
- Strategia inside-out
- Strategia mista

Progettazione concettuale

Nella progettazione **bottom-up**, le **specifiche iniziali** sono suddivise in **componenti** via via piu' piccole, ed in un secondo momento i vari **schemi** sono **integrati** tra loro.



Progettazione concettuale



Progettazione concettuale

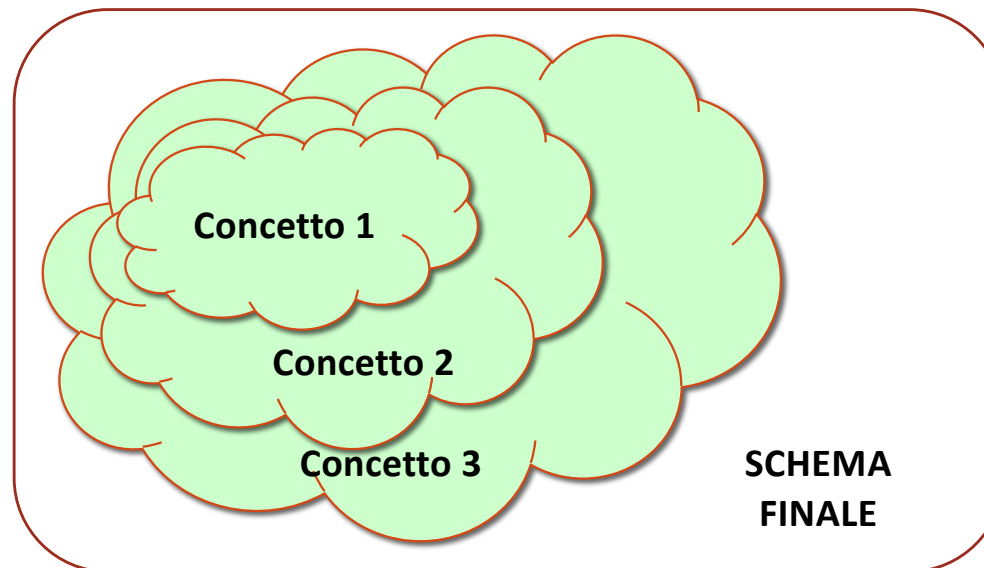


In generale, il documento di specifica potrebbe essere molto complesso e denso di contenuti ... **da dove partire per la costruzione del modello E-R?**

- Strategia top-down
- Strategia bottom-up
- **Strategia inside-out**
- Strategia mista

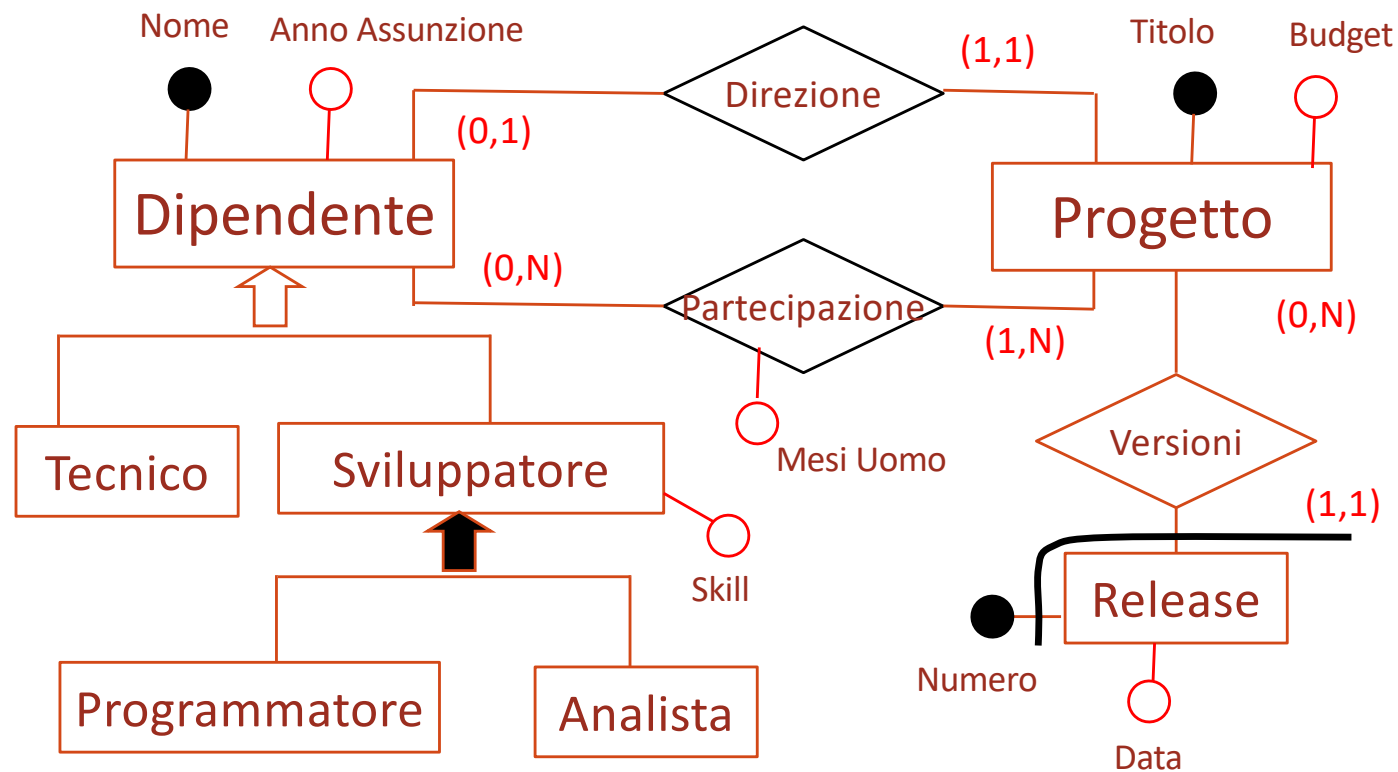
Progettazione concettuale

Nella progettazione **inside-out**, si individuano una serie di concetti importanti e poi si procede a partire da questi verso concetti correlati, con un'estensione a macchia d'olio.



Strategia
inside-out:
caso particolare
della strategia
bottom-up.

Progettazione concettuale



Progettazione concettuale



In generale, il documento di specifica potrebbe essere molto complesso e denso di contenuti ... **da dove partire per la costruzione del modello E-R?**

- Strategia top-down
- Strategia bottom-up
- Strategia inside-out
- **Strategia mista**

Progettazione concettuale

In pratica si utilizza una combinazione delle strategie precedenti, detta anche **strategia mista**.

1. Si individuano i **concetti principali** o più citati
2. Si realizza uno **schema scheletro**
3. **Si decompone** lo schema
4. **Si raffina** lo schema si espande, si integra

In molti casi pratici di una certa complessità, *la strategia mista rappresenta la scelta migliore ...*

Progettazione concettuale

Come garantire tali proprietà?

In generale dipendono dal progettista, ma esistono alcune **metodologie di progettazione concettuale ...**

- Strategie di progettazione
- **Pattern di progettazione**
- Analisi di prestazione

Progettazione concettuale

Non esiste una rappresentazione univoca delle specifiche, nel dubbio è meglio attenersi alle **Regole Concettuali** (RC) del diagramma E-R.

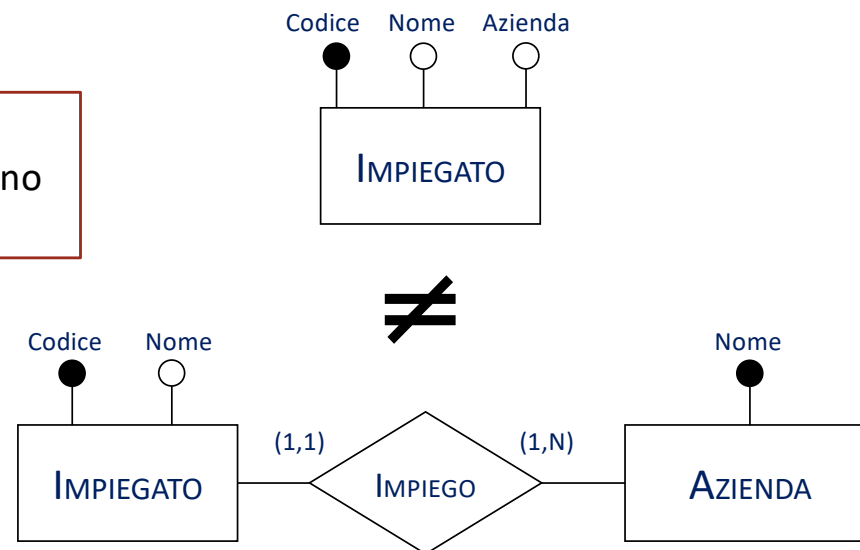
- ✧ **RC1.** Se un concetto ha proprietà significative e descrive oggetti con esistenza autonoma → **Usare Entità**
- ✧ **RC2.** Se un concetto correla due o più entità → **Usare Relazioni**
- ✧ **RC3.** Se un concetto è un caso particolare dell'altro → **Usare Generalizzazioni**

Progettazione concettuale

Esistono molti **pattern** (= **soluzione di problemi ricorrenti**) usati nella progettazione concettuale.

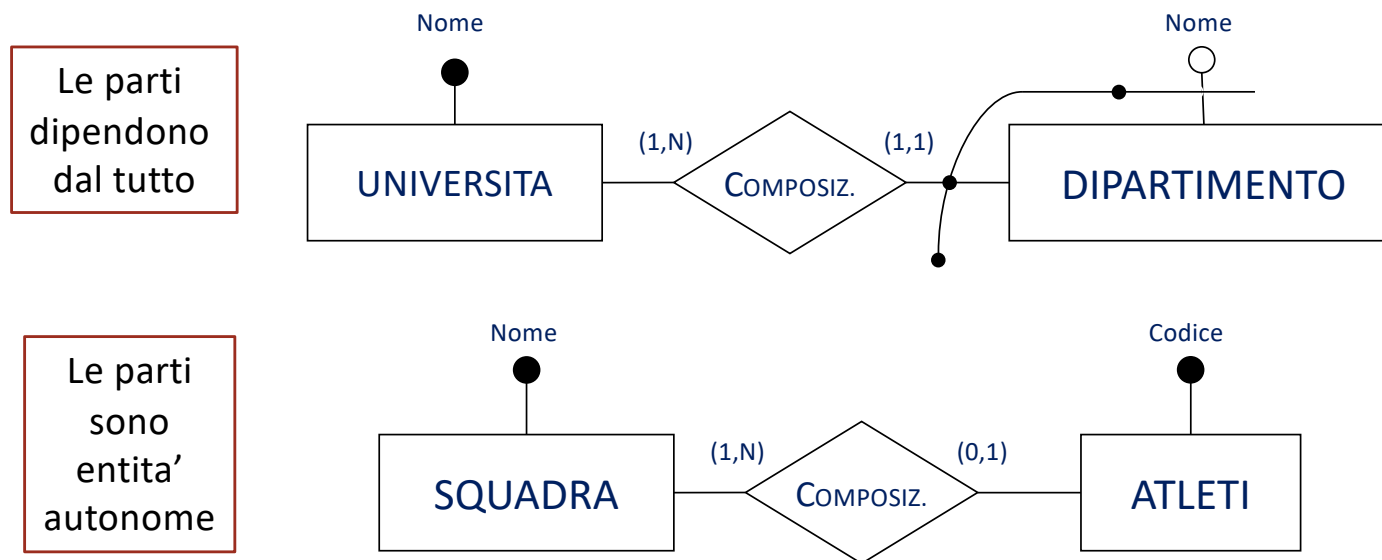
ATTENZIONE!

Questi schemi non sono equivalenti!!



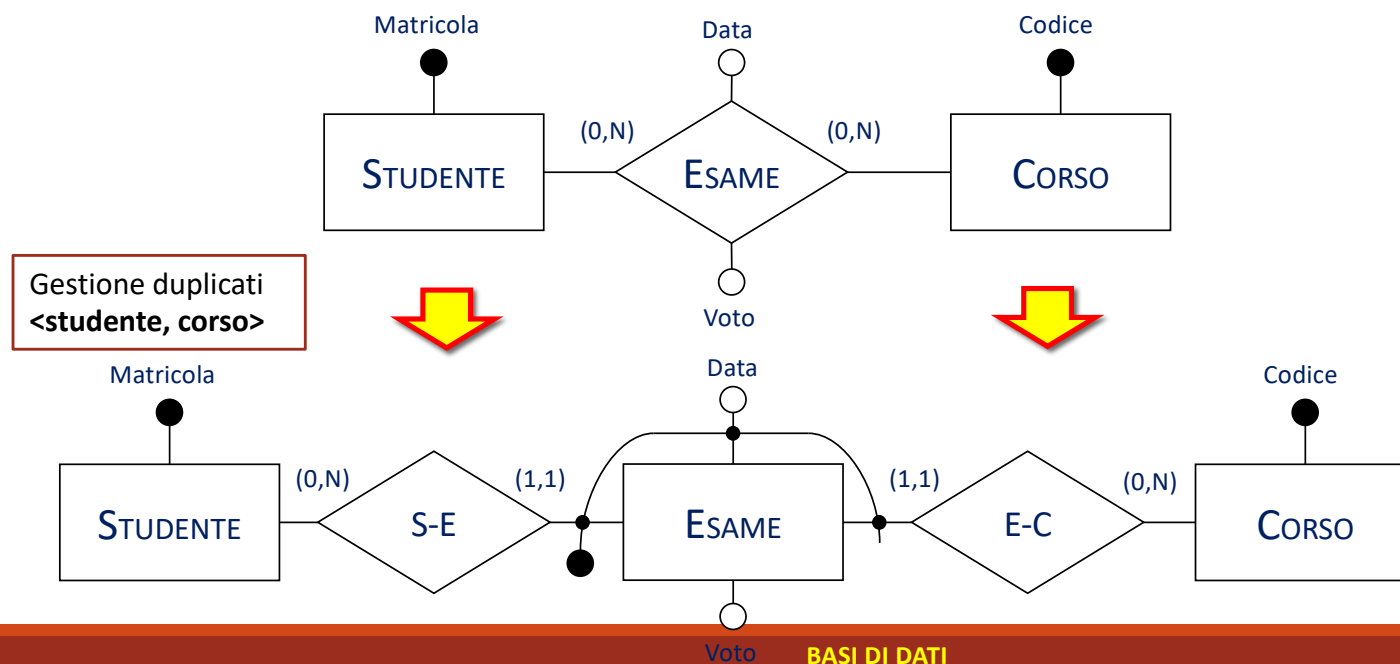
Progettazione concettuale

PATTERN₁. Concetti di tipo “parte-di” attraverso l’utilizzo di **relazioni uno-a-molti**.



Progettazione concettuale

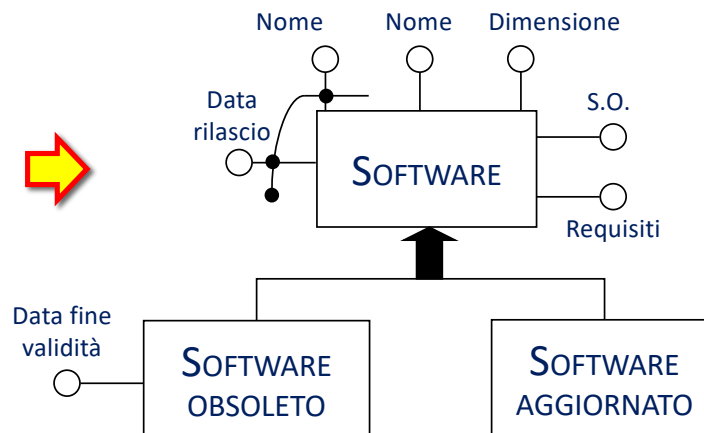
PATTERN₂. Introduzione di nuove entità in relazioni uno-a-molti per la gestione dei duplicati.



Progettazione concettuale

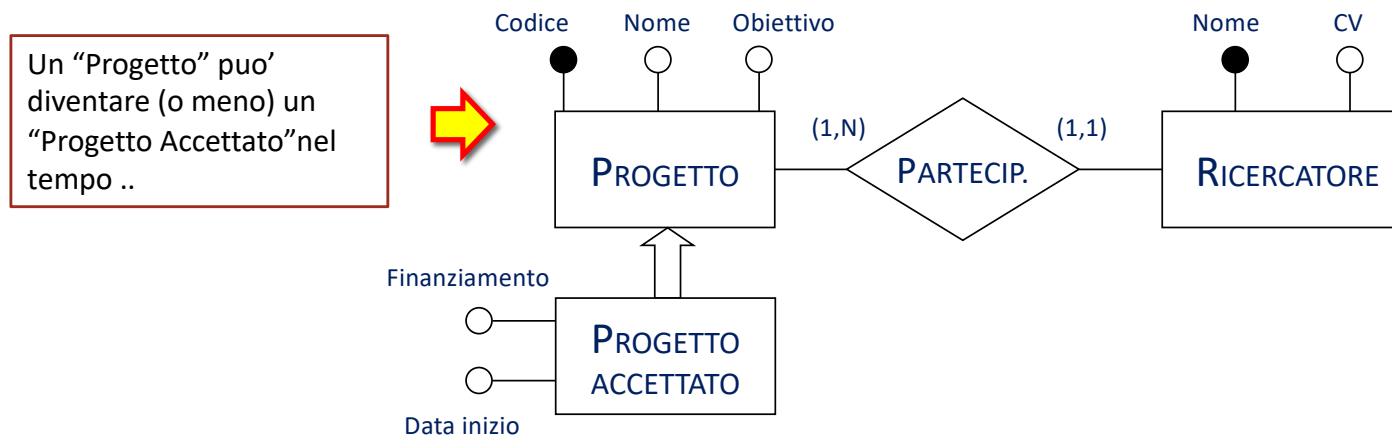
PATTERN₃. Utilizzo di **generalizzazioni** per tenere traccia della **storia di un concetto** (ossia, della sua istanza attuale e di quelle pregresse).

La generalizzazione consente di evitare duplicati tra le entità, inserendo solo gli attributi specifici del concetto aggiornato/storico.



Progettazione concettuale

PATTERN₄. Utilizzo di **generalizzazioni** per tenere traccia dell'**evoluzione nel tempo di un certo concetto** (ossia creazione di nuove istanze diverse dal concetto originario).



Progettazione concettuale

Come garantire tali proprietà?

In generale dipendono dal progettista, ma esistono alcune **metodologie di progettazione concettuale ...**

- Strategie di progettazione
- Pattern di progettazione
- **Analisi di prestazione**

Progettazione concettuale

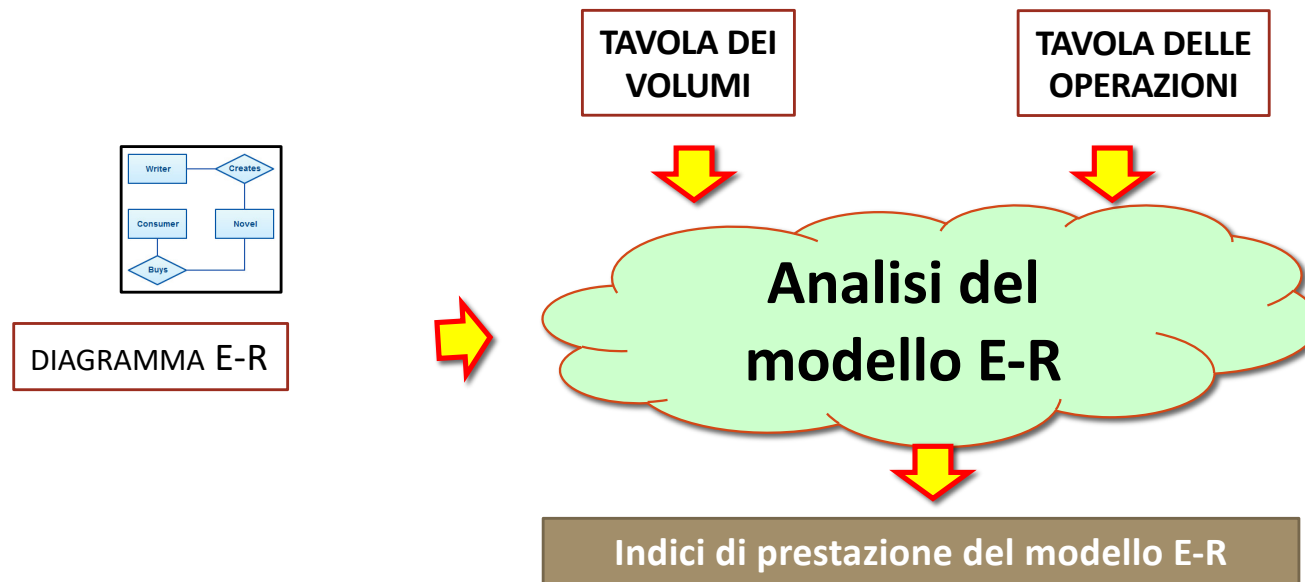
Una volta realizzato il **modello E-R**, è importante analizzarne l'**efficienza dal punto di vista prestazionale**.

Indici di prestazione:

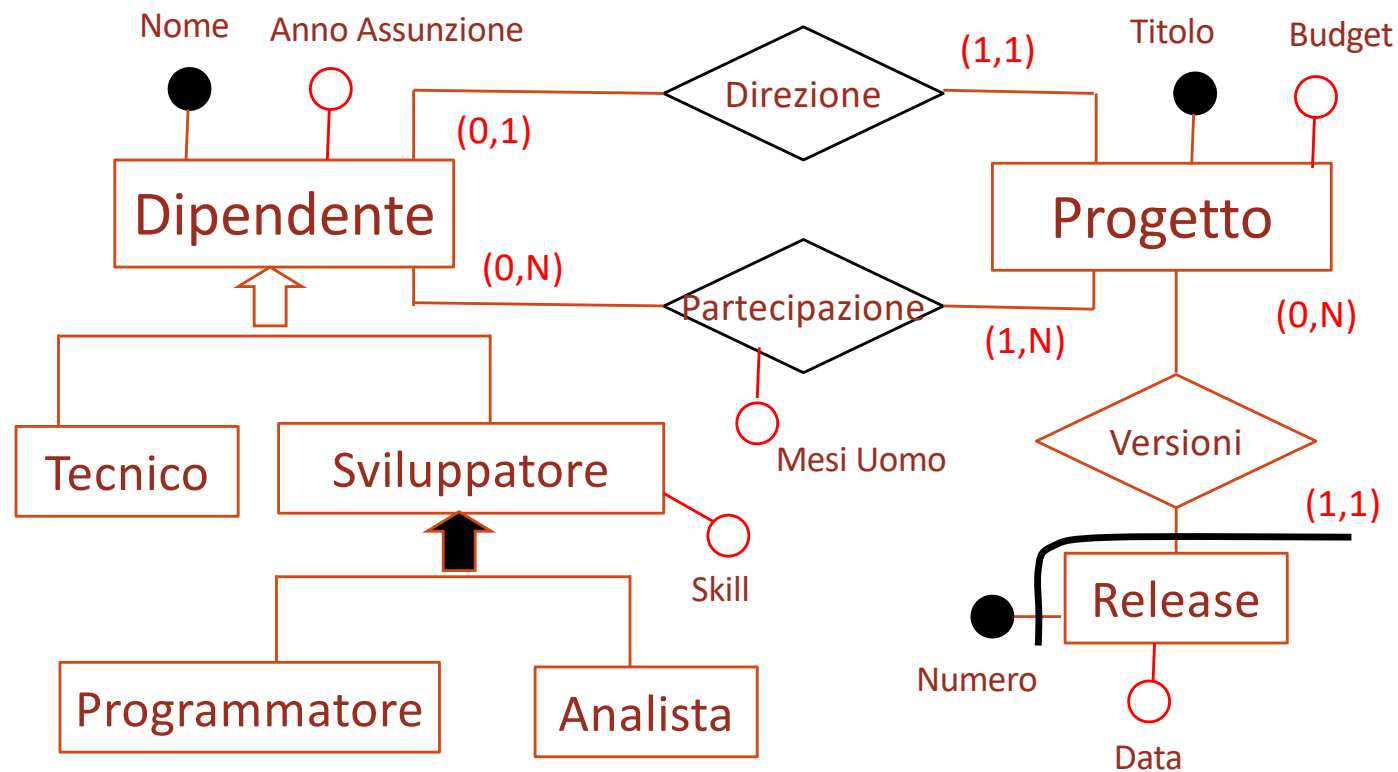
- **Costo operativo** → Numero di entità/associazioni mediamente visitate per implementare una certa operazione sui dati.
- **Occupazione di memoria** → Spazio di memoria necessario per memorizzare i dati.

Progettazione concettuale

Per poter stimare l'**efficienza prestazionale** di uno schema E-R, **abbiamo necessità di informazioni aggiuntive ...**



Progettazione concettuale



Progettazione concettuale

La **tavola dei volumi** fornisce una **stima del numero di occorrenze di entità/ relazioni** presenti nel modello E-R.

Concetto	Tipo	Volume
Progetto	E	100
Release	E	1000
Dipendente	E	500
Tecnico	E	200
Sviluppatore	E	100
Programmatore	E	70
Analista	E	30

Stima del
**numero
medio** di
occorrenze
di una entità

Progettazione concettuale

La **tavola dei volumi** fornisce una **stima del numero di occorrenze di entità/ relazioni** presenti nel modello E-R.

Concetto	Tipo	Volume
Direzione	R	100
Partecipazione	R	2000
Versioni	R	1000

Stima
del **numero
medio** di
occorrenze
di una relazione

- **Assunzione1:** Ogni progetto ha in media 10 release.
- **Assunzione2:** Ad ogni progetto lavorano in media 20 dipendenti.

Progettazione concettuale

La **tavola delle operazioni** definisce:

- L'**insieme delle operazioni** che devono essere implementate.
- La **tipologia** delle operazioni (interattive/batch).
- La **frequenza** delle operazioni (es. 100 al giorno).

D. Chi fornisce le informazioni delle tavole?

R. Spesso la **raccolta ed analisi dei requisiti** ...

Progettazione concettuale

- ✧ **Operazione₁**: assegnare un dipendente ad un progetto.
- ✧ **Operazione₂**: visualizzare tutti i dati di un un progetto, della release associate e del direttore.
- ✧ **Operazione₃**: per ciascun progetto, visualizzare tutti i dati dei dipendenti associati.

Operazione	Tipo	Frequenza
Operazione ₁	I	10 volte/giorno
Operazione ₂	I	100 volte/giorno
Operazione ₃	B	5 volte/giorno

Progettazione concettuale

Data un'operazione O di tipo T , definiamo il suo **costo** $c(O_T)$ come:

$$c(O_T) = f(O_T) \cdot w_T \cdot (\alpha \cdot NC_{write} + NC_{read})$$

$f(O_T)$ → Frequenza dell'operazione

NC_{read} → Numero di **accessi in lettura** a componenti (entita'/relazioni) dello schema

NC_{write} → Numero di **accessi in scrittura** a componenti (entita'/relazioni) dello schema

w_T → **Peso** dell'operazione (interattiva/batch)

α → **Coefficiente moltiplicativo** delle operazioni in scrittura

Progettazione concettuale

- **Operazione₁**: assegnare un dipendente ad un progetto (frequenza: 10 volte al gg).

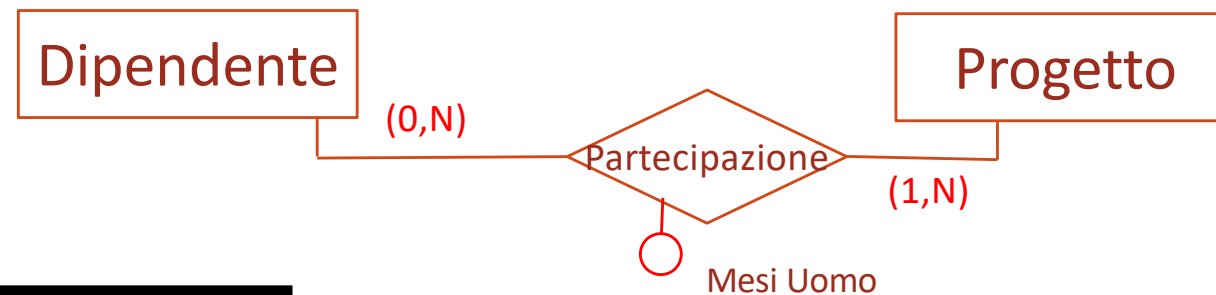


TAVOLA DEGLI ACCESSI

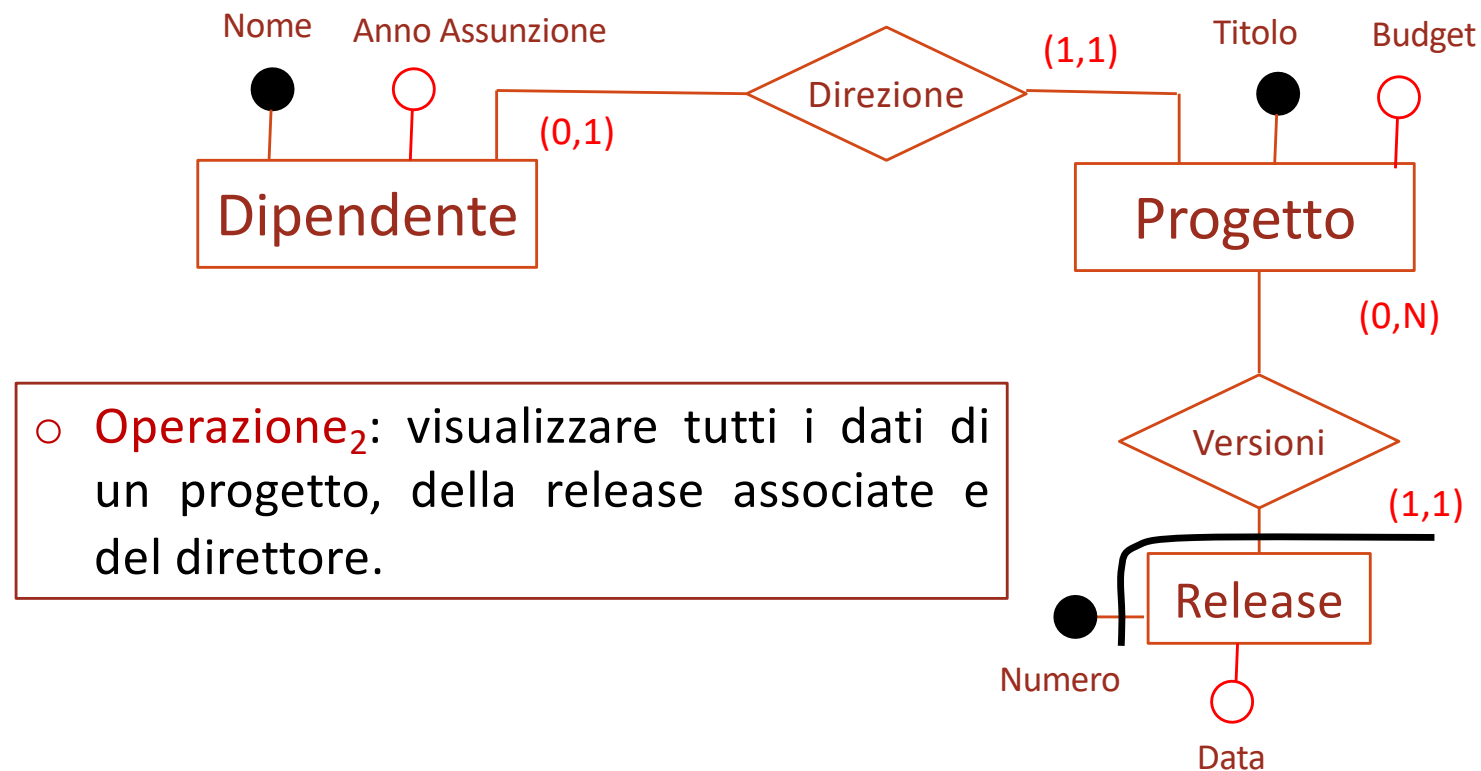
Concetto	Costrutto	Accessi	Tipo
Partecipazione	Relazione	1	W

Parametri:

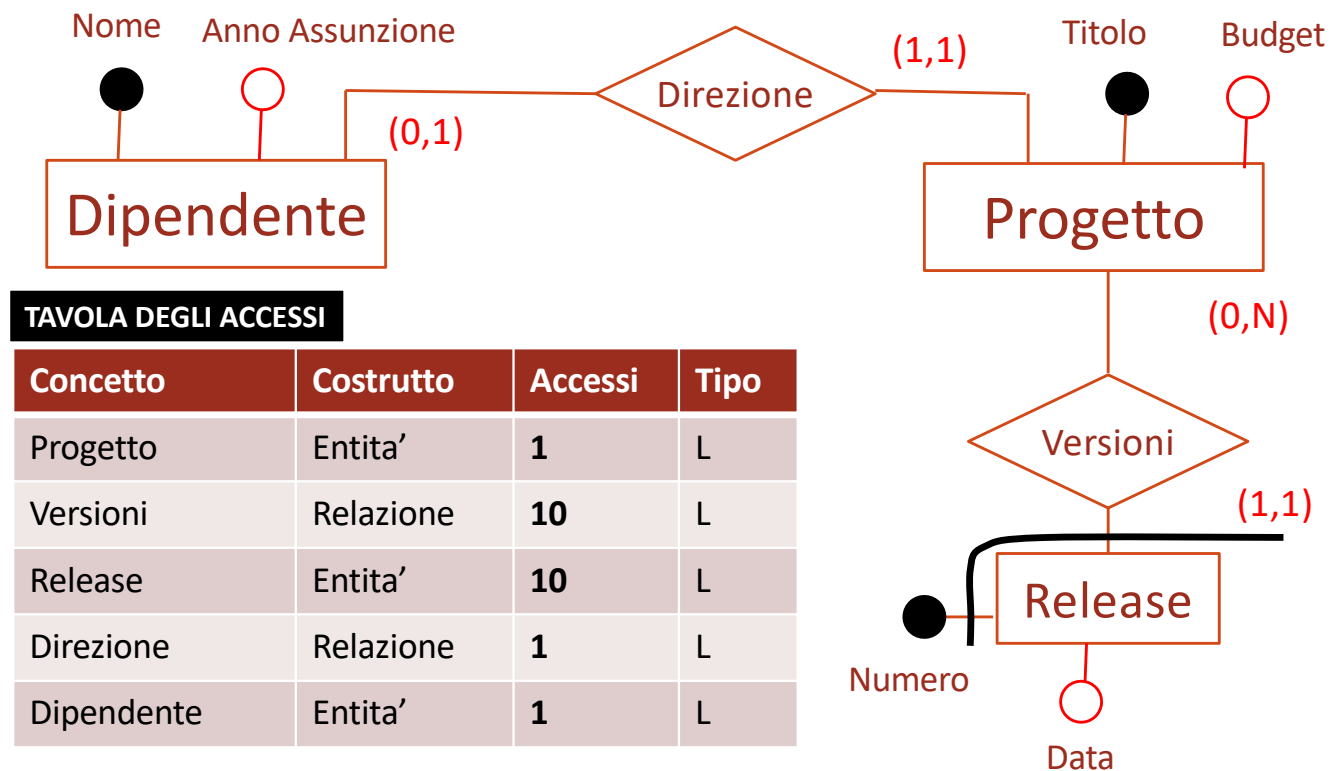
$$w_I = 0.5, w_B = 0.5 \quad \alpha = 2$$

$$C(Op1) = 10 \cdot 0.5 \cdot (2 \cdot 1 + 0) = 10$$

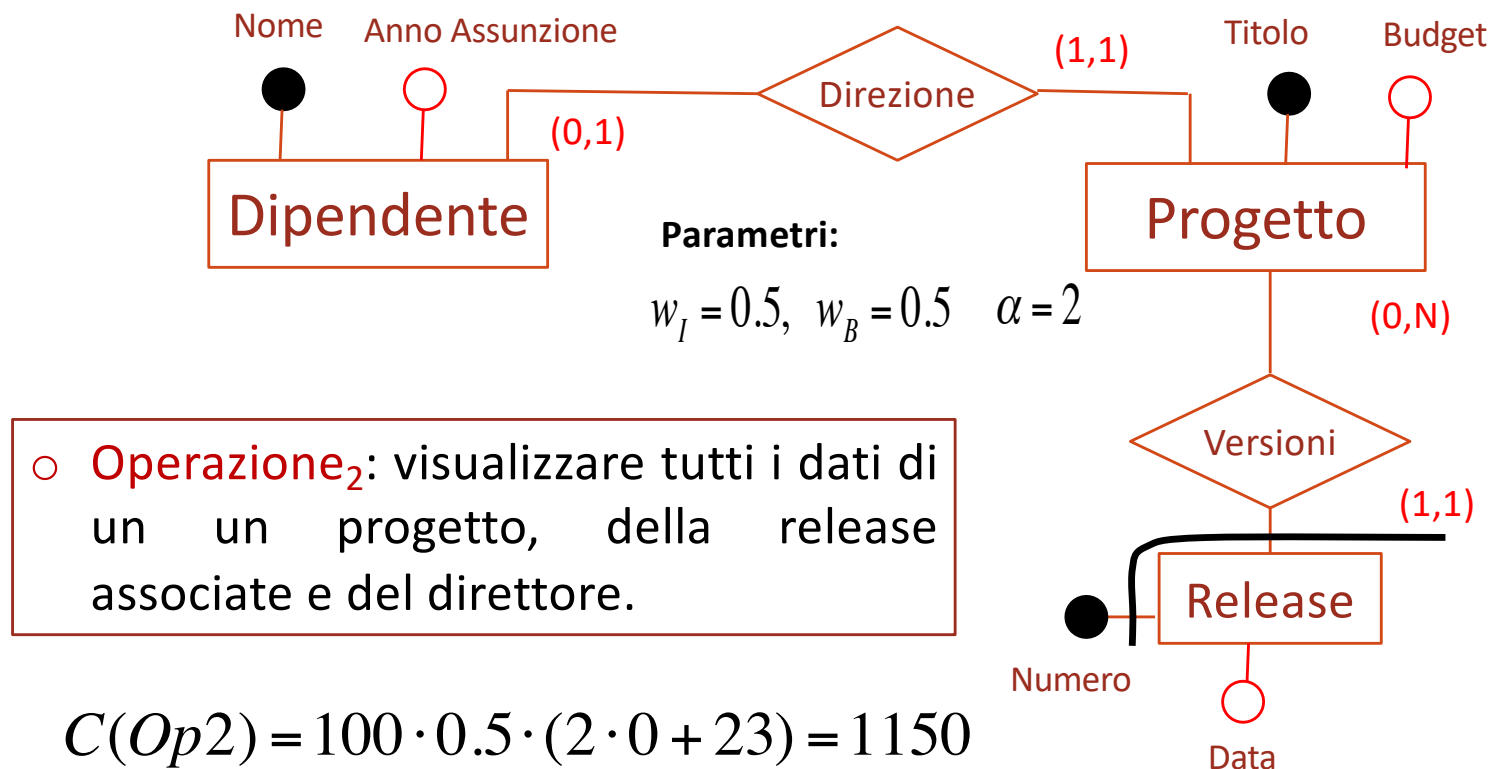
Progettazione concettuale



Progettazione concettuale



Progettazione concettuale



Progettazione concettuale

- **Operazione₃**: per ciascun progetto, visualizzare tutti i dati del progetto e dei dipendenti associati.

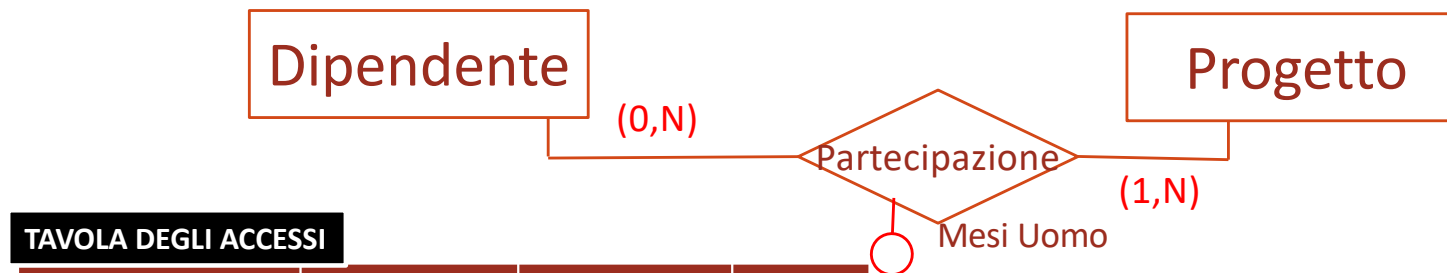


TAVOLA DEGLI ACCESSI

Concetto	Costrutto	Accessi	Tipo
Progetto	Entita'	100	R
Partecipazione	Relazione	2000	R
Dipendente	Entita'	2000	R

Parametri:

$$w_I = 0.5, w_B = 0.5 \quad \alpha = 2$$

$$C(Op3) = 5 \cdot 0.5 \cdot (0 + 4100) = 10250$$

Progettazione concettuale

Dato uno schema S , ed un'insieme di operazioni sui dati O^1, O^2, \dots, O^n , con costi $c(O^1), c(O^2), c(O^n)$, il **costo dello schema** e' definito come:

$$c(S) = \sum_{i=1}^n c(O^i)$$

Nel nostro esempio:

$c(\text{Op1}) \rightarrow 10$	+
$c(\text{Op2}) \rightarrow 1150$	+
$c(\text{Op3}) \rightarrow 10250$	

$c(S) \rightarrow 11410$

Progettazione concettuale

Dato uno schema S , ed un'insieme di operazioni sui dati $O^1, O^2, \dots O^n$, con costi $c(O^1), c(O^2), c(O^n)$, il **costo dello schema** e' definito come:

$$c(S) = \sum_{i=1}^n c(O^i)$$

Obiettivo del progettista è quello di determinare lo schema E-R **di costo minimo ...**

Q1. Come determinare lo **schema minimale**?

Q2. Cosa dire dell'**occupazione di memoria**?

Progettazione concettuale

Conoscendo la tavola dei volumi, il tipo di ciascun attributo e la sua dimensione su disco, è possibile stimare l'**occupazione di memoria dello schema ...**

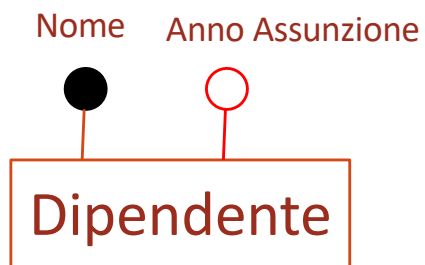
$$M(S) = \sum_{\text{entità } e} V(e) \cdot \text{size}(e) + \sum_{\text{relazione } r} V(r) \cdot \text{size}(r)$$

$V(e), \text{size}(e) \rightarrow$ Tabella dei volumi e dimensione in termini di occupazione di memoria dell'entità e .

$V(r), \text{size}(r) \rightarrow$ Tabella dei volumi e dimensione in termini di occupazione di memoria della relazione r .

Progettazione concettuale

ESEMPIO: Come stimare l'occupazione di memoria dell'entità "Dipendente" ...



Concetto	Tipo	Volume
Dipendente	E	500

Attributo	Tipo	Dimensione
Nome	Stringa (max 8 caratteri)	1B*8=8B
AnnoAss.	Intero	4B

Memoria Occupata (*Dipendente*)=500*(8B+4B)=6000B

Progettazione concettuale

In pratica: si cerca di determinare il miglior **trade-off** tra **occupazione di memoria e costo delle operazioni dello schema ...**

Gli indici di prestazione di un diagramma E-R sono forniti come input alla fase di **progettazione logica**, e sono utilizzati per:

- Traduzione del modello concettuale
- Analisi delle ridondanze.

Progettazione concettuale

Ricapitolando:

- STEP 0: **Analisi dei requisiti**
- STEP1: **Progettazione Concettuale**
 - ✧ STEP 1.1: **Diagramma E-R**
 - ✧ STEP 1.2: **Dizionario** delle entità
 - ✧ STEP 1.2: **Dizionario** delle relazioni
 - ✧ STEP 1.3: Tabella delle **business rules**
 - ✧ STEP 1.4: **Analisi indici di stima**