



Progettazione Concettuale di Basi di Dati

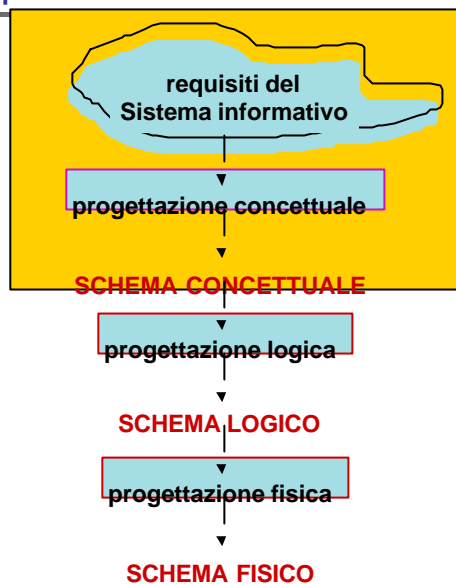
Sistemi Informativi L

Corso di Laurea in Ingegneria dei Processi Gestionali
A.A. 2003/2004

Docente: Prof. Wilma Penzo



Il primo passo...





Raccolta dei requisiti

- I requisiti devono innanzitutto essere **acquisiti**
- Le fonti possono essere molto diversificate tra loro:
 - **utenti**, attraverso
 - interviste
 - documentazione apposita
 - **documentazione esistente**
 - normative (leggi, regolamenti di settore)
 - regolamenti interni, procedure aziendali
 - realizzazioni preesistenti
 - **modulistica**
- La raccolta dei requisiti è un'attività difficile e non standardizzabile; in genere procede di pari passo con la fase di analisi (la prima analisi stimola nuove domande, ecc.)



Interagire con gli utenti

- È un'attività da considerare con molta attenzione, in quanto:
 - **utenti diversi** possono fornire **informazioni diverse**
 - utenti a livello più alto hanno spesso una **visione più ampia** ma **meno dettagliata**
- In generale, risulta utile:
 - effettuare spesso **verifiche** di comprensione e coerenza
 - verificare anche per mezzo di **esempi** (generali e relativi a casi limite)
 - richiedere **definizioni** e **classificazioni**
 - far evidenziare gli **aspetti essenziali** rispetto a quelli marginali



Requisiti: documentazione descrittiva

- Regole generali:
 - scegliere il corretto **livello di astrazione**
 - standardizzare la struttura delle frasi
 - suddividere le frasi articolate
 - separare le frasi sui **dati** da quelle sulle **funzioni** (operazioni)

- Per meglio evidenziare i concetti che sono espressi nei requisiti, è opportuno:
 - costruire un **glossario dei termini**
 - individuare **omonimi** e **sinonimi** e unificare i termini
 - rendere esplicito il **riferimento fra termini**
 - riorganizzare le frasi per concetti



Requisiti: un esempio

Base di dati bibliografica

Si vogliono organizzare i dati di interesse per automatizzare la gestione dei riferimenti bibliografici



Requisiti BD bibliografica (1)

Si vogliono organizzare i dati di interesse per automatizzare la gestione dei riferimenti bibliografici, **con tutte le informazioni da riportarsi in una bibliografia.**

Per ogni pubblicazione deve esistere un codice identificante costituito da sette caratteri, indicanti le iniziali degli autori, l'anno di pubblicazione e un carattere aggiuntivo per la discriminazione delle collisioni.



Requisiti BD bibliografica (2)

Si vogliono organizzare i dati di interesse per automatizzare la gestione dei riferimenti bibliografici, con tutte le informazioni da riportarsi in una bibliografia.

Le pubblicazioni sono di due tipi, monografie (per le quali interessano editore, data e luogo di pubblicazione) e articoli su rivista (con nome della rivista, volume, numero, pagine e anno di pubblicazione); per entrambi i tipi si debbono ovviamente riportare i nomi degli autori.

Per ogni pubblicazione deve esistere un codice identificante

.....



Un altro esempio più articolato

Società di formazione

Si vuole realizzare una base di dati per una società che eroga corsi, di cui vogliamo rappresentare i dati dei **partecipanti** ai corsi e dei **docenti**.

Per gli **studenti** (circa 5000), identificati da un codice, si vuole memorizzare il codice fiscale, il cognome, l'età, il sesso, il **luogo** di nascita, il nome dei loro attuali datori di lavoro, i **posti** dove hanno lavorato in precedenza insieme al periodo, l'indirizzo e il numero di telefono, i **corsi** che hanno frequentato (i corsi sono in tutto circa 200) e il **giudizio finale**.

Rappresentiamo anche i **seminari** che stanno attualmente frequentando e, per ogni giorno, i **luoghi** e le ore dove sono tenute le lezioni.



Società di formazione (2)

I corsi hanno un codice, un **titolo** e possono avere varie edizioni con date di inizio e fine e numero di partecipanti.

Se gli studenti sono liberi professionisti, vogliamo conoscere l'area di interesse e, se lo possiedono, il **titolo**. Per quelli che **lavorano alle dipendenze di altri**, vogliamo conoscere invece il loro livello e la posizione ricoperta.

Per gli **insegnanti** (circa 300), rappresentiamo il cognome, l'età, il **posto** dove sono nati, il nome del corso che insegnano, quelli che hanno insegnato nel passato e quelli che possono insegnare. Rappresentiamo anche tutti i loro recapiti telefonici. I docenti possono essere dipendenti interni della società o collaboratori esterni.



Glossario dei termini, omonimi e sinonimi

- Raramente i requisiti espressi in linguaggio naturale sono privi di ambiguità. È infatti frequente il caso di
 - Omonimi:** lo stesso termine viene usato per descrivere concetti differenti (es: libro e copia di libro, posto: di lavoro e geografico)
 - Sinonimi:** termini diversi vengono usati per descrivere lo stesso concetto (es: studente e partecipante)
- Un modo conveniente per rappresentare sinteticamente i concetti più rilevanti emersi dall'analisi è il **glossario dei termini**, il cui scopo è fornire per ogni concetto rilevante:
 - Una breve descrizione del concetto
 - Eventuali sinonimi
 - Relazioni con altri concetti del glossario stesso



Glossario dei termini: esempio

Termine	Descrizione	Sinonimi	Collegamenti
Partecipante	Persona che partecipa ai corsi. Può essere un dipendente o un professionista	Studente	Corso, Datore
Docente	Docente dei corsi. Può essere un collaboratore esterno	Insegnante	Corso
Corso	Corso organizzato dalla società. Può avere più edizioni	Seminario	Docente, Partecipante
Datore	Datori di lavoro attuali o passati dei partecipanti ai corsi	Posti	Partecipante



Ristrutturazione dei requisiti

- Oltre a costruire il glossario, per semplificare le analisi successive, è utile riformulare i requisiti:
 - Eliminare le omonimie
 - Usare un termine univoco per ogni concetto
 - Riorganizzare le frasi raggruppandole in base al concetto cui si riferiscono

Nell'esempio:

- Frasi di carattere generale
- Frasi riferite ai partecipanti
- Frasi riferite ai docenti
- Frasi riferite ai corsi
- Frasi riferite ai datori di lavoro



Esempio: frasi relative ai partecipanti

Prima:

Per gli **studenti** (circa 5000), identificati da un codice, si vuole memorizzare il codice fiscale, il cognome, l'età, il sesso, il **luogo** di nascita, il nome dei loro attuali datori di lavoro, i **posti** dove hanno lavorato in precedenza insieme al periodo, l'indirizzo e il numero di telefono, i **corsi** che hanno frequentato (i corsi sono in tutto circa 200) e il **giudizio finale**. Rappresentiamo anche i **seminari** che stanno attualmente frequentando... []. Se gli studenti sono liberi professionisti, vogliamo conoscere l'area di interesse e, se lo possiedono, il **titolo**. Per quelli che **lavorano alle dipendenze di altri**, vogliamo conoscere invece il loro livello e la posizione ricoperta.

Dopo:

Per i **partecipanti** (circa 5000), identificati da un codice, si vuole memorizzare il codice fiscale, il cognome, l'età, il sesso, la **città** di nascita, il nome dei loro attuali datori di lavoro e di **quelli precedenti** (insieme alle date di inizio e fine rapporto), compresi indirizzo e numero di telefono, le **edizioni dei corsi** che stanno attualmente frequentando e quelli che hanno frequentato nel passato, con la relativa **votazione finale in decimi**. Per i liberi professionisti si memorizzi l'area di interesse e, se lo possiedono, il **titolo professionale**. Per i **dipendenti** si memorizzino il livello e la posizione ricoperta.



Dai concetti allo schema E/R

- Va sempre ricordato che un concetto non è di per sé un'entità, un'associazione, un attributo, o altro

DIPENDE DAL CONTESTO!

- Come regole guida, un concetto verrà rappresentato come
 - Entità
 - se ha proprietà significative e descrive oggetti con esistenza autonoma
 - Attributo
 - se è semplice e non ha proprietà
 - Associazione
 - se correla due o più concetti
 - Generalizzazione/specializzazione
 - se è caso più generale/particolare di un altro



Strategie di progettazione

- Per affrontare progetti complessi è opportuno adottare uno specifico modo di procedere, ovvero una **strategia di progettazione**
- I casi notevoli sono:
 - Strategia **top-down**
 - Si parte da uno schema iniziale molto astratto ma completo, che viene successivamente raffinato fino ad arrivare allo schema finale
 - Strategia **bottom-up**
 - Si suddividono le specifiche in modo da sviluppare semplici schemi parziali ma dettagliati, che poi vengono integrati tra loro
 - Strategia **inside-out**
 - Lo schema si sviluppa "a macchia d'olio", partendo dai concetti più importanti, aggiungendo quelli ad essi correlati, e così via



Strategia top-down

specifiche



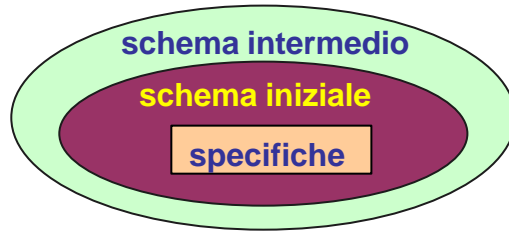
Strategia top-down

schema iniziale

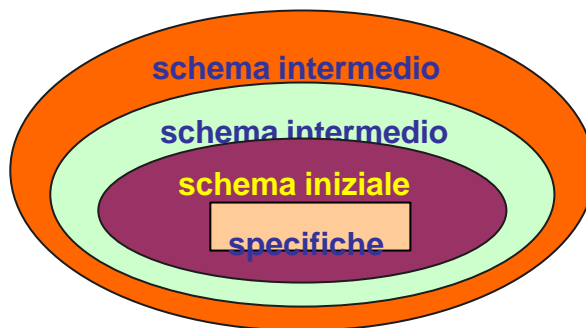
specifiche

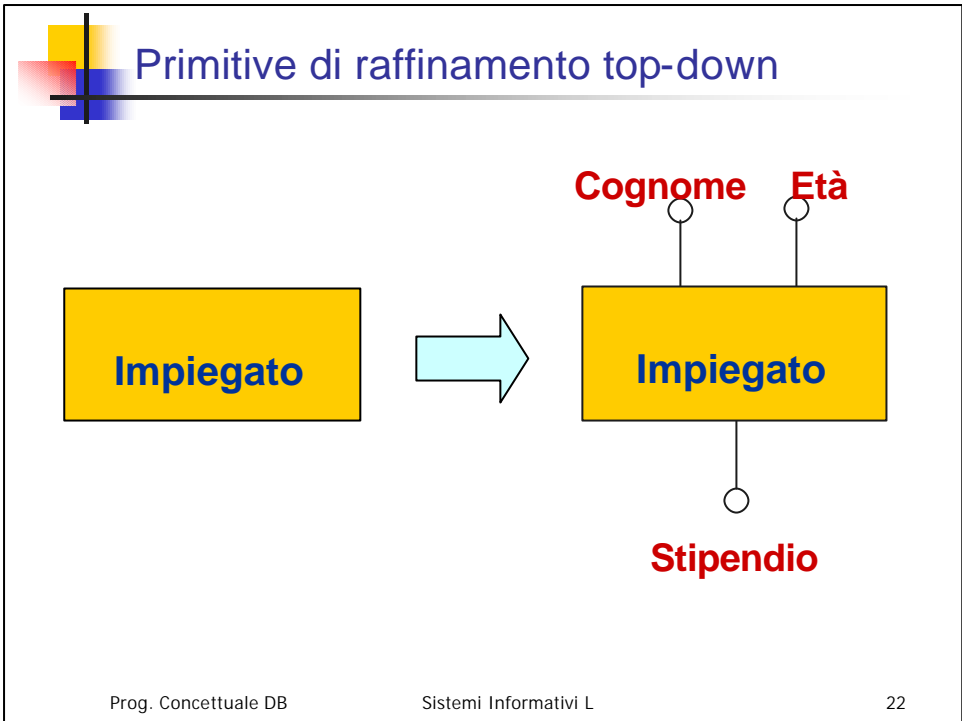
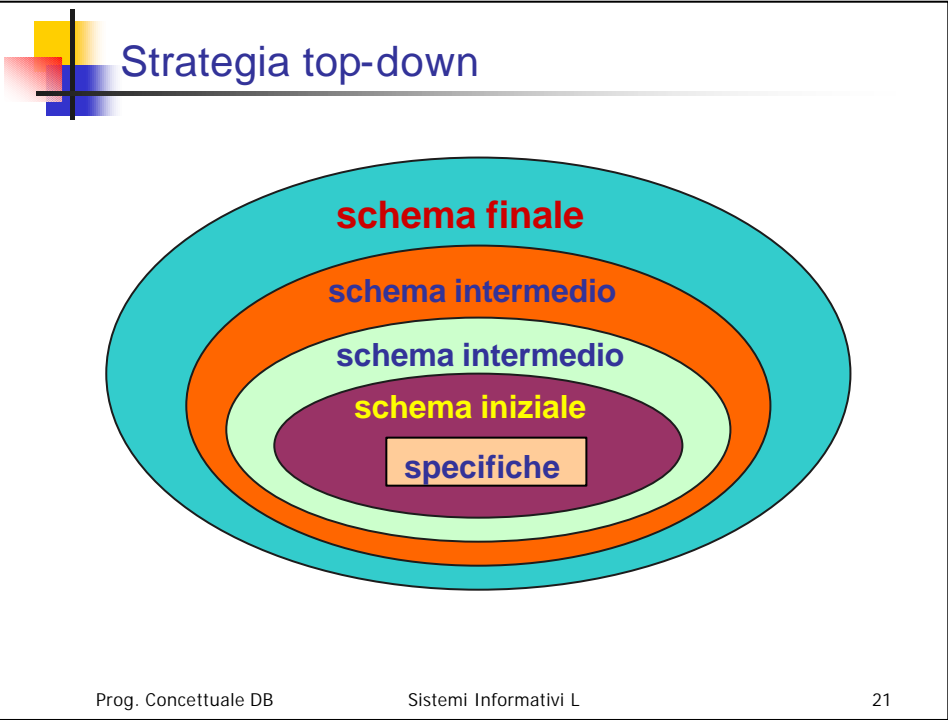


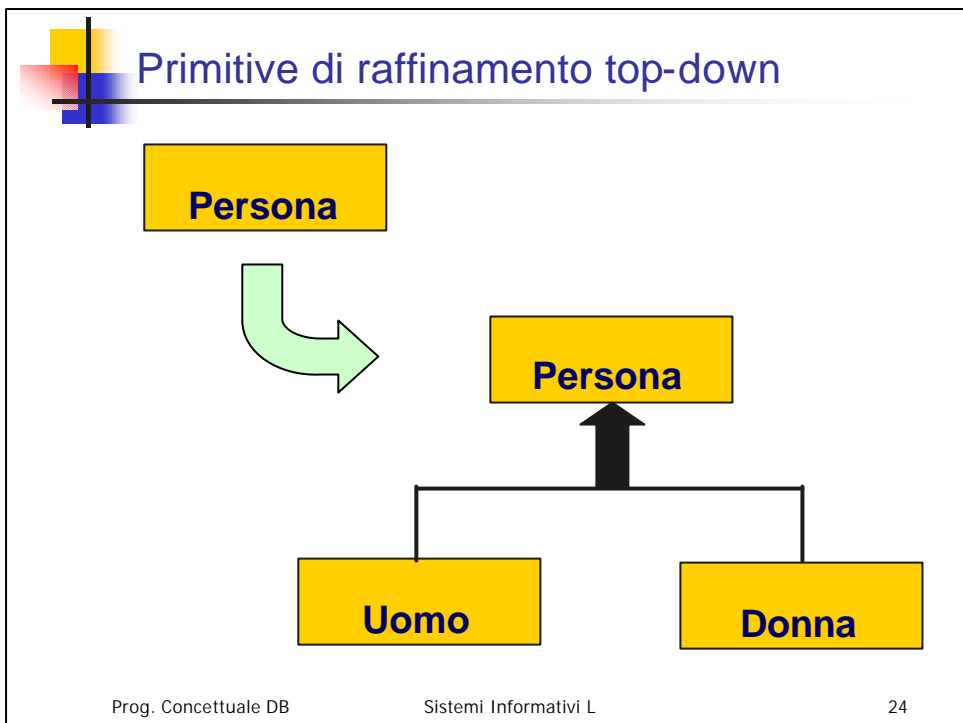
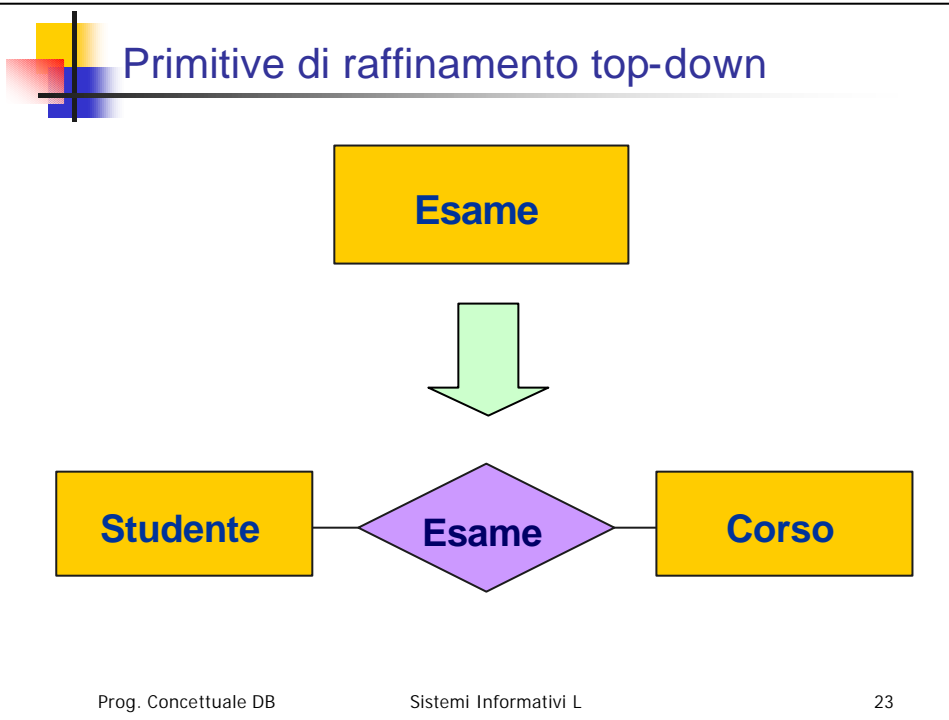
Strategia top-down

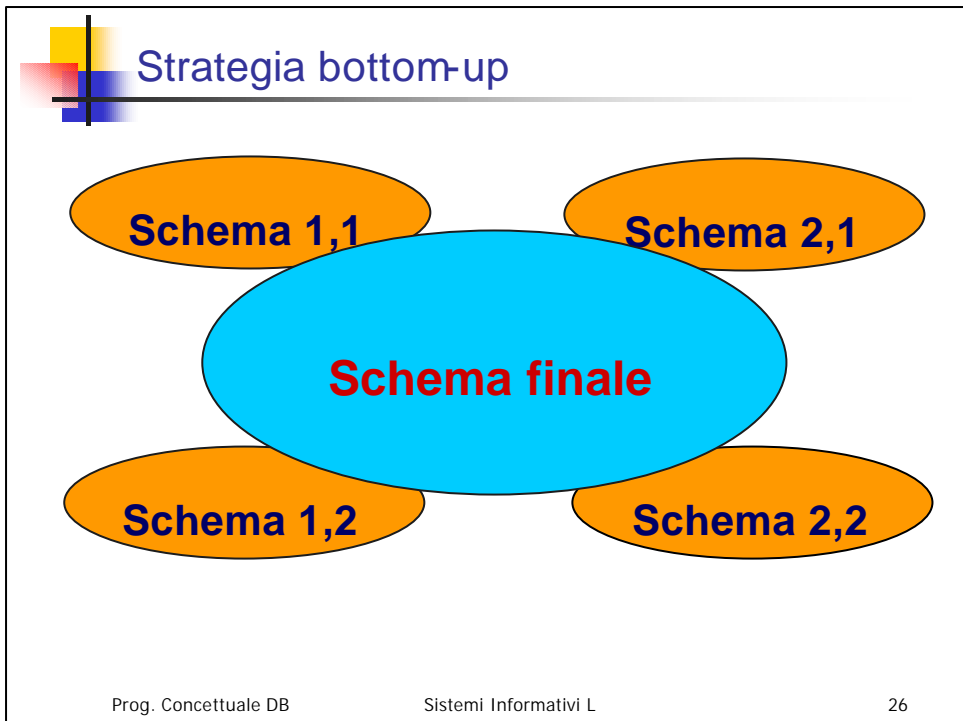
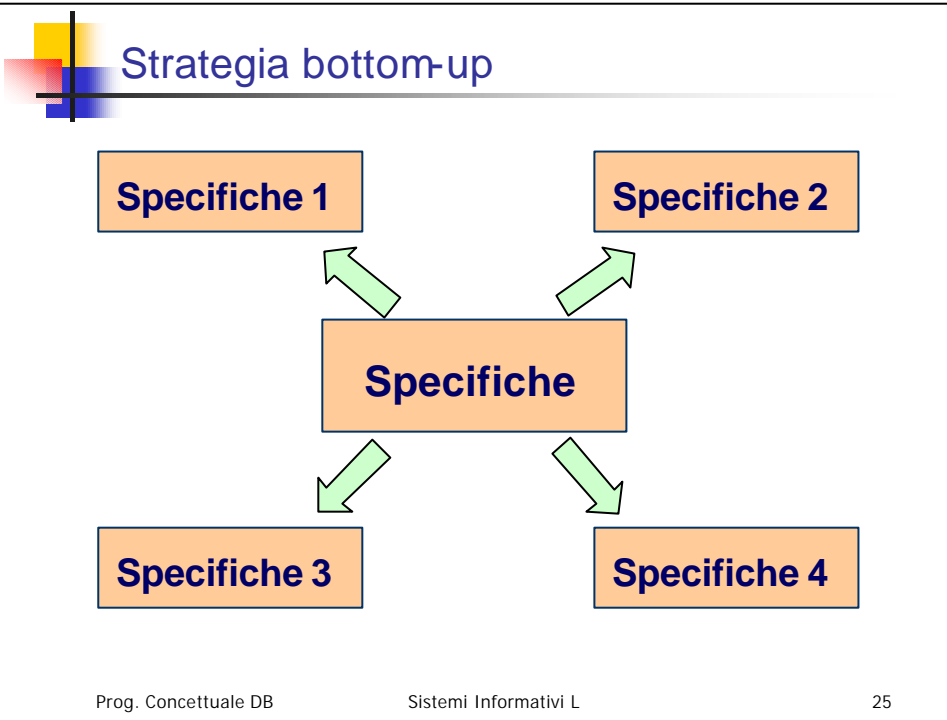


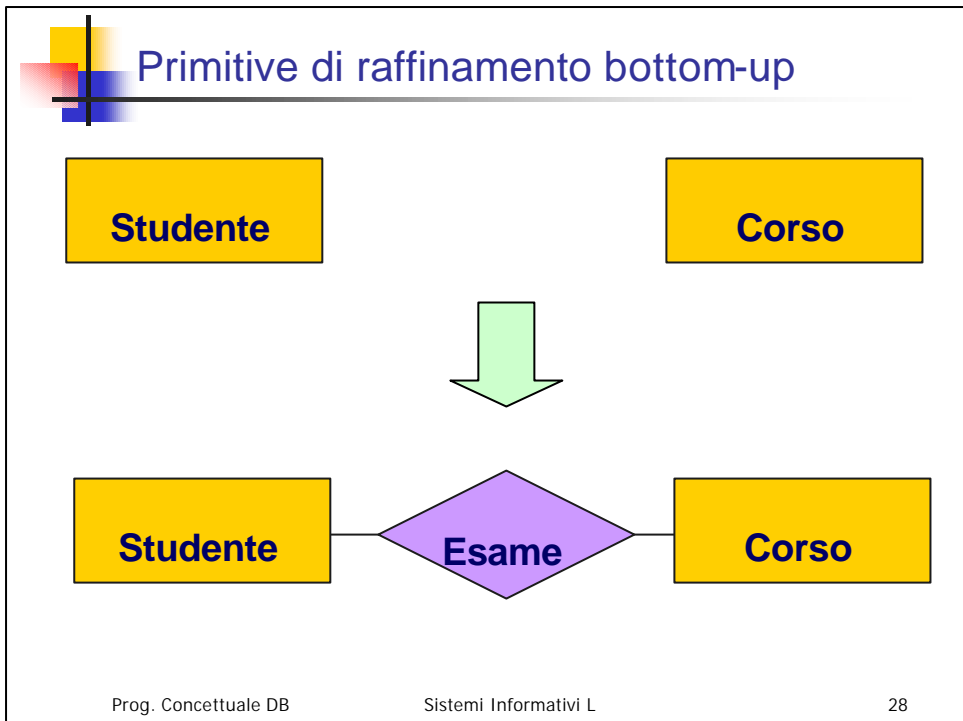
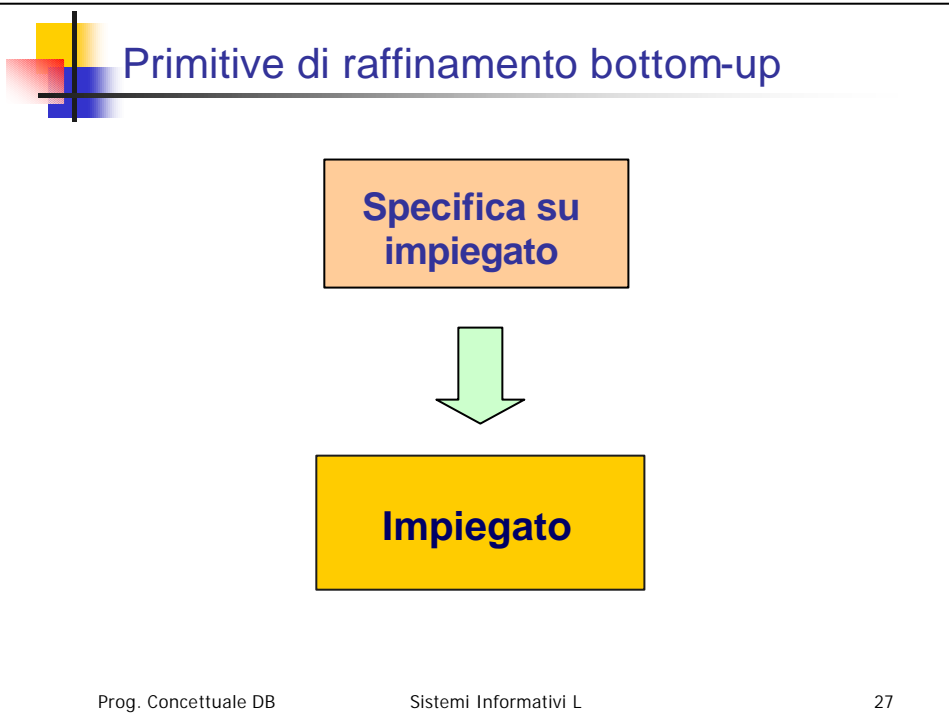
Strategia top-down





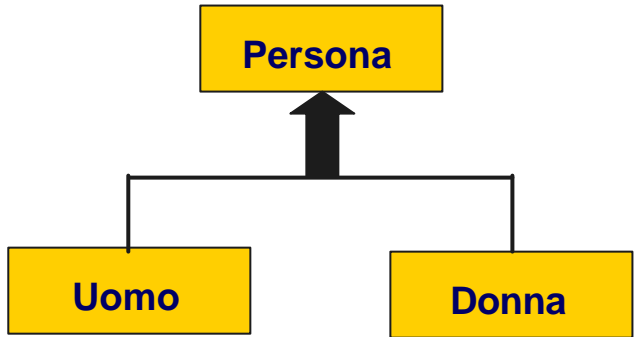




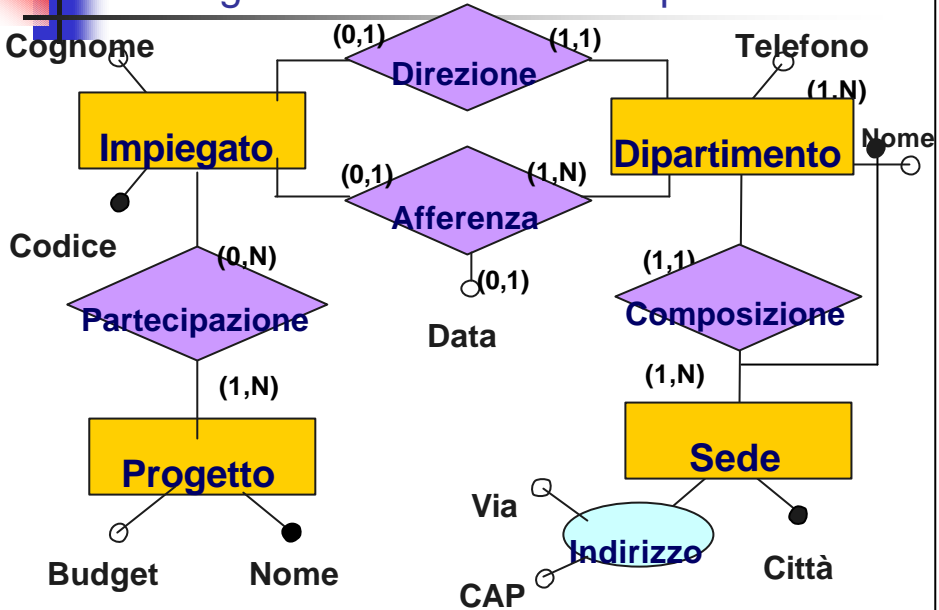




Primitive di raffinamento bottom-up



Strategia inside-out: un esempio





Strategie: pro e contro

Top-down

- + non è inizialmente necessario specificare i dettagli
- richiede sin dall'inizio una visione globale del problema, non sempre ottenibile in casi complessi

Bottom-up

- + permette una ripartizione delle attività
- richiede una fase di integrazione

Inside-out

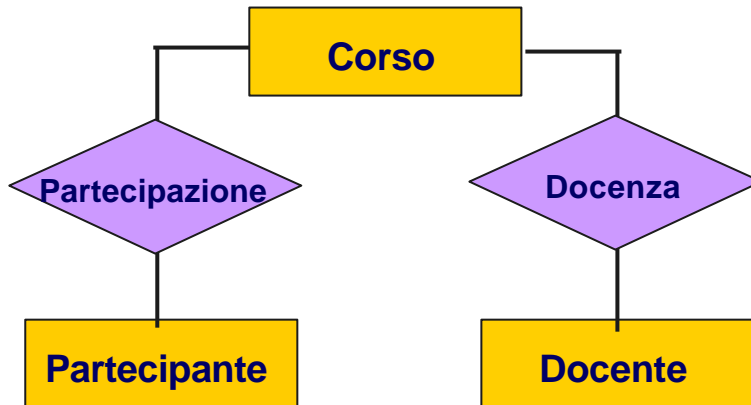
- + non richiede passi di integrazione
- richiede ad ogni passo di esaminare tutte le specifiche per trovare i concetti non ancora rappresentati.



Un approccio "misto"

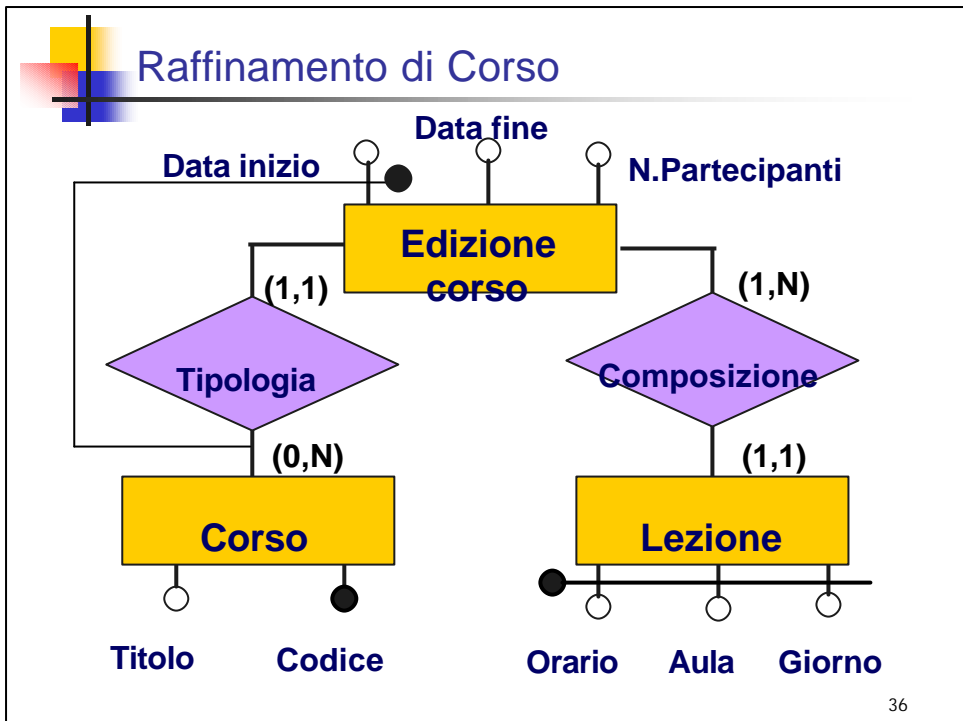
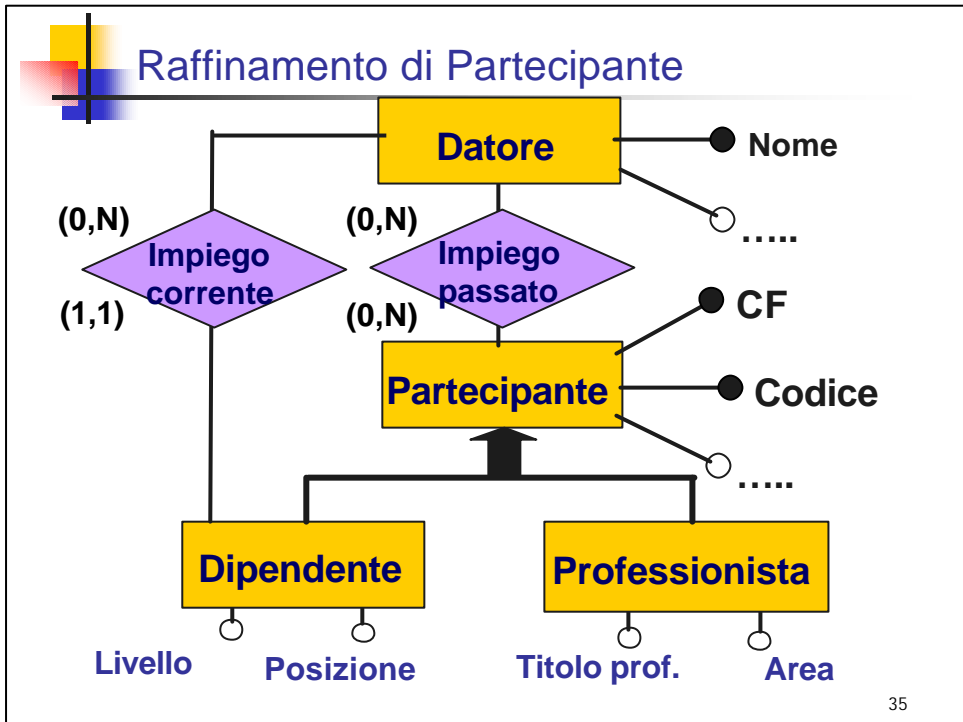
- Nella pratica si fa spesso uso di una strategia ibrida, nella quale:
 - 1 si individuano i concetti principali e si realizza uno **schema scheletro**, che contiene solamente i concetti più importanti
 - 2 sulla base di questo si può decomporre
 - 3 poi si raffina, si espande, si integra
- ... vediamo cosa succede nel caso della società di formazione...

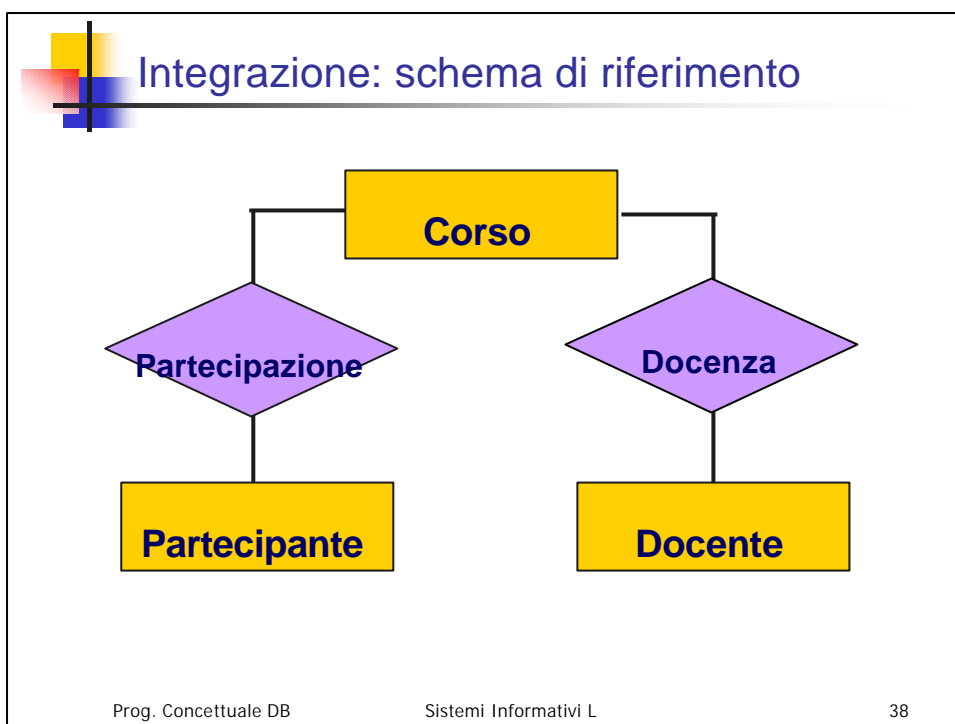
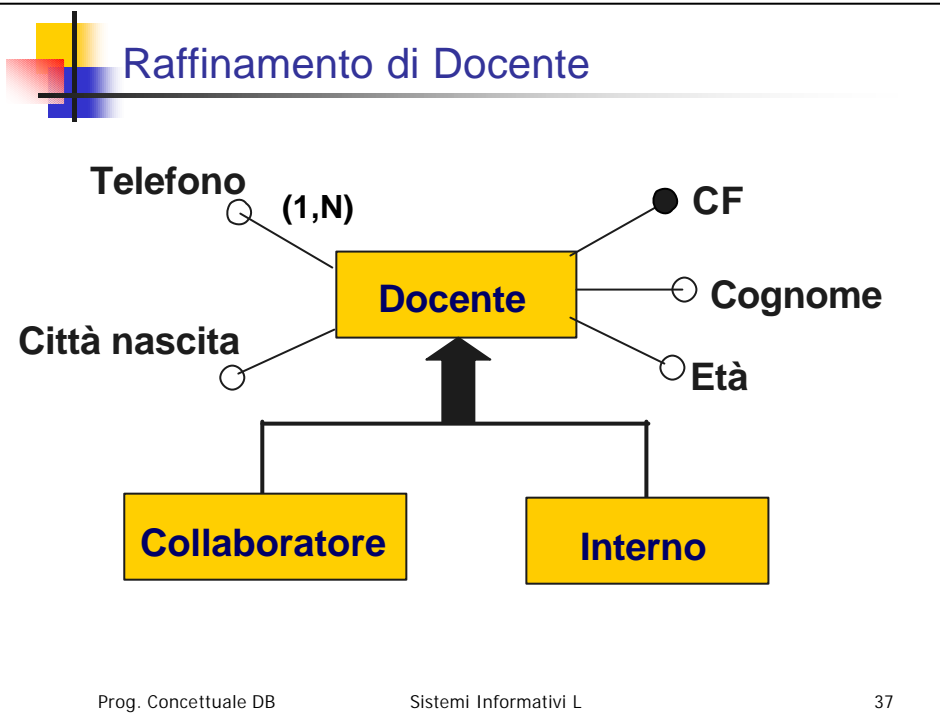
Società di formazione: schema scheletro

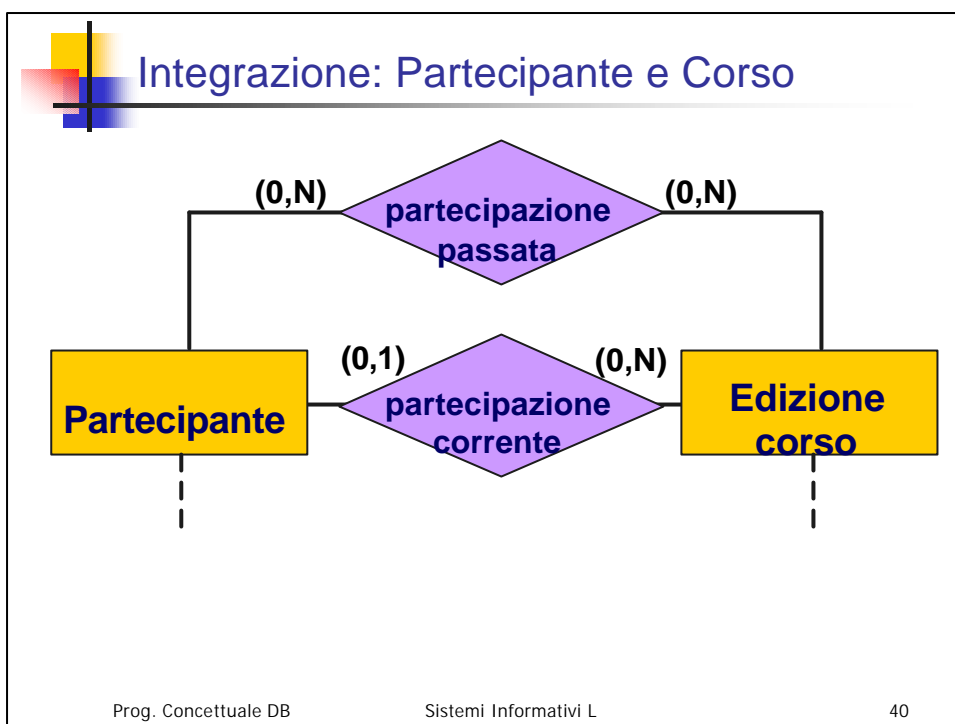
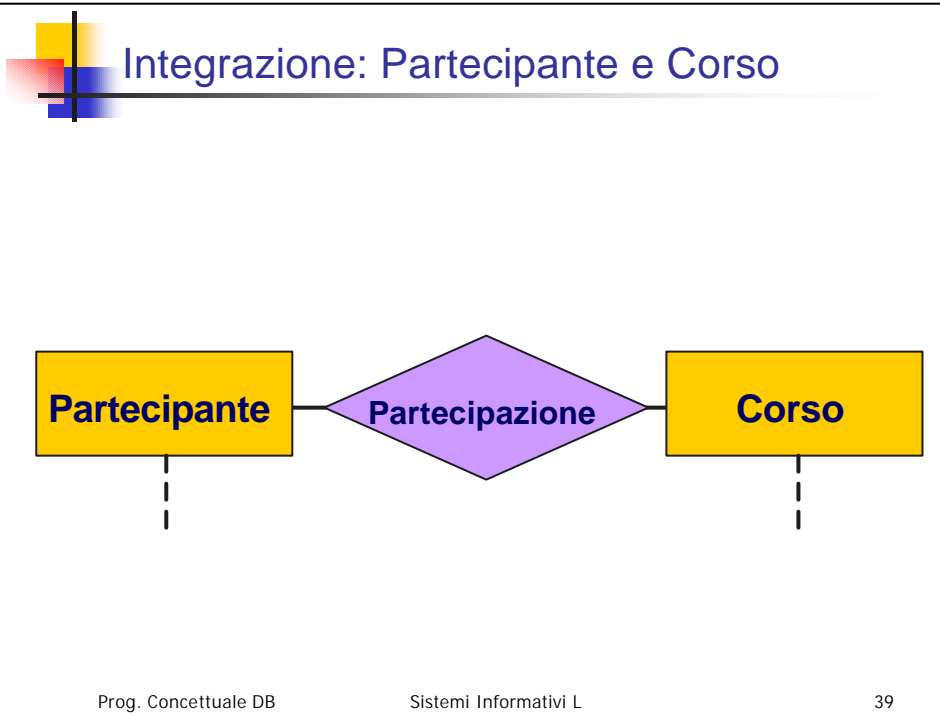


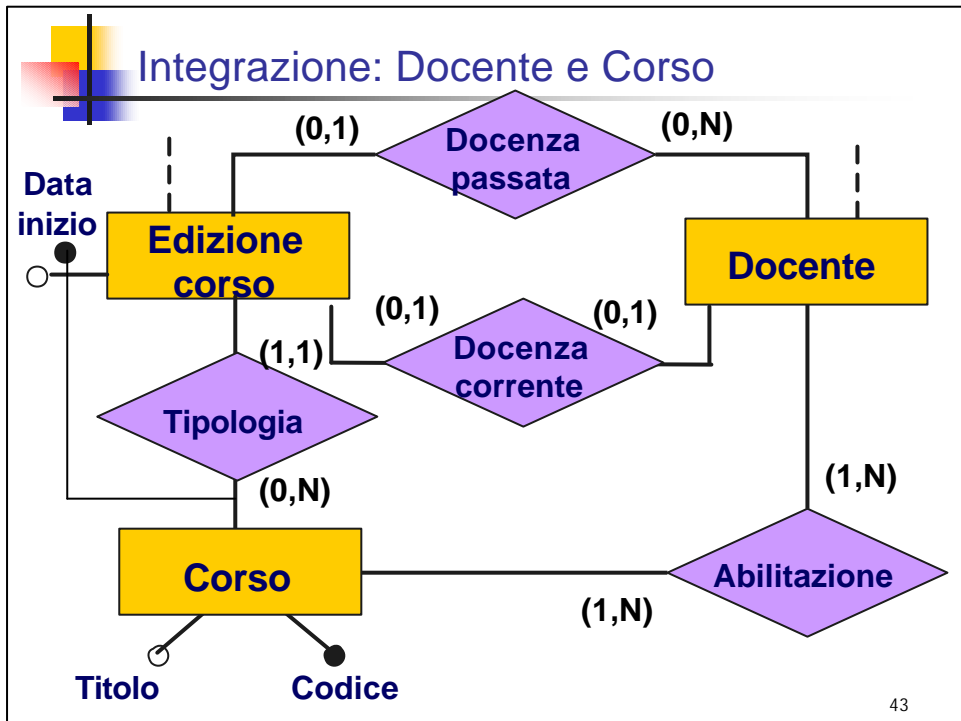
Frase relative ai partecipanti

Per i partecipanti (circa 5000), identificati da un codice, rappresentiamo il codice fiscale, il cognome, l'età, il sesso, la città di nascita, il nome dei loro attuali datori di lavoro e di quelli precedenti (insieme alle date di inizio e fine rapporto), compresi indirizzo e numero di telefono, le edizioni dei corsi che stanno attualmente frequentando e quelli che hanno frequentato nel passato, con la relativa votazione finale in decimi. Per i liberi professionisti di memorizziamo l'area di interesse e, se lo possiedono, il titolo professionale. Per i dipendenti si memorizzino il livello e la posizione ricoperta.









43

Qualità di uno schema concettuale

- Lo schema E/R deve essere verificato accuratamente per verificare che risponda a requisiti di:
 - **Correttezza**
 - Non devono essere presenti errori (sintattici o semantici)
 - **Completezza**
 - Tutti i dati di interesse sono specificati
 - **Leggibilità**
 - Riguarda anche aspetti prettamente estetici dello schema
 - **Minimalità**
 - È importante capire se esistono elementi ridondanti nello schema; in alcuni casi ciò non è un problema, ma può essere viceversa una scelta di progettazione volta a favorire l'esecuzione di alcune operazioni



Metodologia basata sulla strategia mista

Analisi dei requisiti

- Analizzare i requisiti ed eliminare le ambiguità
- Costruire un glossario dei termini, raggruppare i requisiti

Passo base

- Definire uno schema scheletro con i concetti più rilevanti

Passo di decomposizione (se necessario o appropriato)

- decomporre i requisiti con riferimento ai concetti nello schemascheletro

Passo iterativo (da ripetere finché non si è soddisfatti)

- Raffinare i concetti presenti sulla base delle loro specifiche
- Aggiungere concetti per descrivere specifiche non descritte

Passo di integrazione (se si è decomposto)

- integrare i vari sottoschemi in uno schema complessivo, facendo riferimento allo schema scheletro

Analisi di qualità (ripetuta e distribuita)

- Verificare le qualità dello schema e modificarlo



Conclusioni

- La **fase di analisi dei requisiti** è fondamentale per poter progettare una base di dati che rispetti i requisiti
- Mancando la possibilità di standardizzarla, tale fase si avvale necessariamente di **regole di buon senso** e di una serie di **strumenti** che riducono il rischio di commettere errori grossolani, oltre a costituire una valida documentazione
- Per la **progettazione dello schema E/R** sono possibili diverse **strategie**, di cui quella **mista** è senz'altro la più diffusa e adeguata anche nel caso di progetti estremamente complessi