

Basi di dati

Concetti introduttivi

ESEMPIO

INSEGNAMENTI
Fisica, Analisi,
Informatica



Aule



Docenti



Studenti

- Entità
- Relazioni
- Interrogazioni

Database

Tabella (I)

STUDENTE

Attributi

Nome	Data di Nascita	Indirizzo	Matricola
Luca Neri	27/10/1980	Via Cavour 5, Milano	123456
Mario Rossi	4/7/1981	Via Roma 4, Pisa	123768
Paolo Bianchi	12/1/1983	Via Reno 2, Bologna	123987

Campo Record

Un record definisce una *relazione* tra i valori dei suoi campi

Tabella (II)

- **DOMINIO** di un attributo definisce l'insieme dei valori che un campo relativo a quell'attributo può assumere
 - *Testo, Valore Numerico, Data, ...*
- **SCHEMA** di una tabella è costituito dal nome della tabella seguito dal nome e dal dominio dei suoi attributi
 - *Studente {Nome: Testo; Data di Nascita: Data; Indirizzo: Testo; Matricola: Numero}*
- **ISTANZA** di una tabella all'istante t è l'insieme dei record che sono presenti nella tabella a tale istante

Notazione: schema

Indicheremo lo schema di un DB secondo le seguenti notazioni alternative

Studente {Nome, Data di Nascita, Indirizzo, Matricola}

Studente
Matricola
Nome
Indirizzo
Data di Nascita

Notazione: istanza

Indicheremo l'istanza di un DB ad un certo istante con la seguente notazione

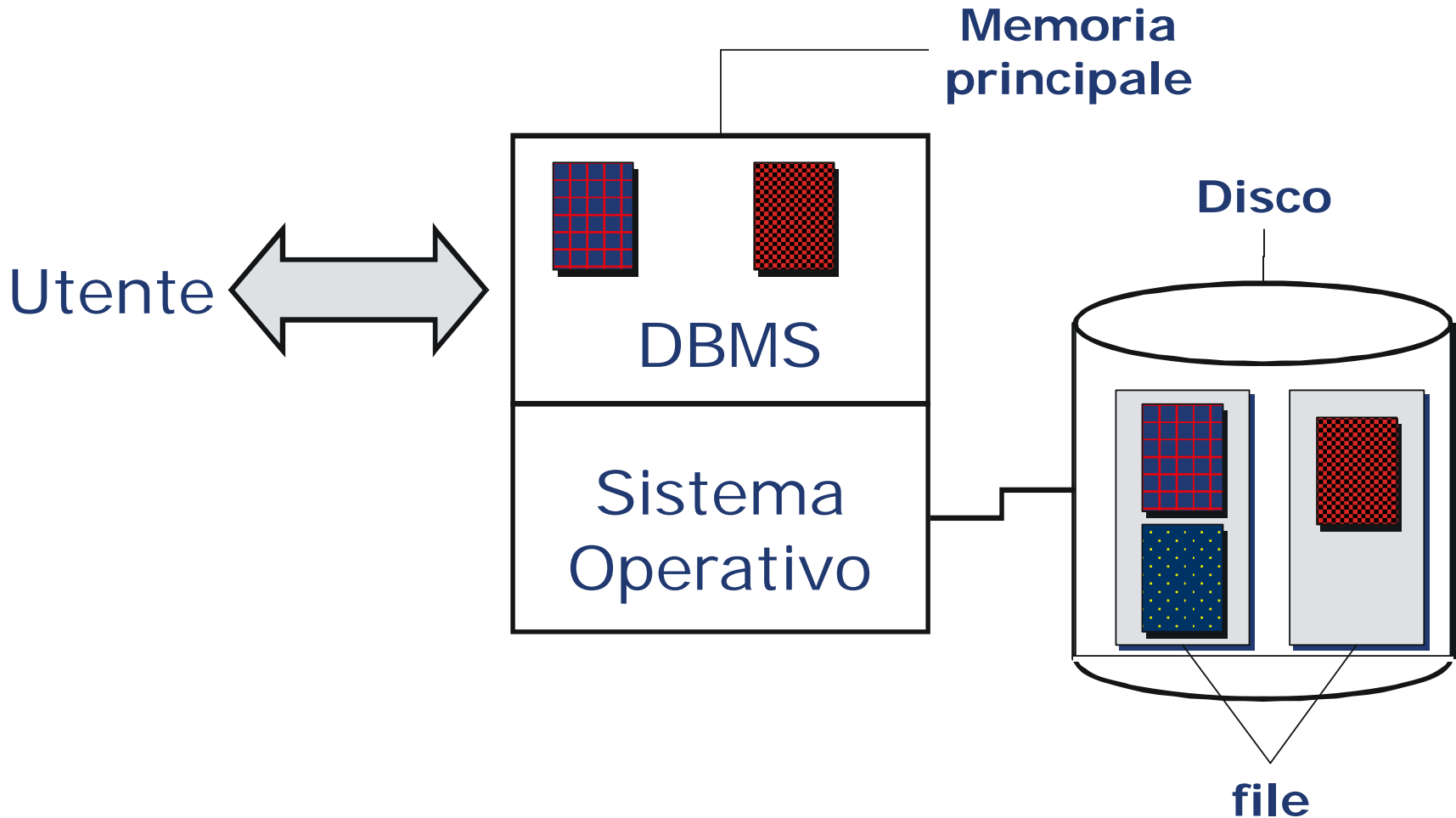
STUDENTE

Nome	Data di Nascita	Indirizzo	Matricola
Luca Neri	27/10/1980	Via Cavour 5, Milano	123456
Mario Rossi	4/7/1981	Via Roma 4, Pisa	123768
Paolo Bianchi	12/1/1983	Via Reno 2, Bologna	123987
:	:	:	:

Database Relazionale

- Un database relazionale è una collezione di tabelle
- Lo schema di un database è costituito dall'elenco dei nomi delle tabelle presenti nel database
- L'istanza del database all'istante t è costituito dall'insieme delle istanze a tale istante di tutte le tabelle presenti nel database

Database Management System



DBMS

Un DBMS è in grado di gestire collezioni di dati che siano

- **Grandi.** Un DB ha dimensioni maggiori della memoria primaria e quindi è memorizzato in memoria secondaria (un DB può avere dimensioni dell'ordine di 1000 GB)
- **Condivise.** Applicazioni ed utenti diversi debbono poter accedere, secondo modalità opportune, a dati comuni
- **Persistenti.** Hanno un tempo di vita indipendente dai programmi che li usano

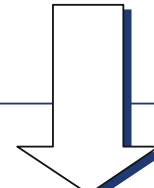
DBMS

Un DBMS deve garantire

- **Affidabilità**, cioè la capacità del sistema di conservare intatto il contenuto del DB in caso di malfunzionamenti hw e sw (backup, recovery)
- **Privatezza**, cioè ogni utente viene abilitato a svolgere solo determinate azioni sui dati
- **Efficienza**, cioè la capacità di svolgere le operazioni utilizzando un insieme di risorse (tempo, spazio) che sia accettabile per gli utenti

Chiave Primaria (una prima definizione)

La *chiave primaria* di una tabella è un attributo tale che non possono esistere due record distinti che hanno lo stesso valore nei campi relativi a quell'attributo



Nome	Data di Nascita	Indirizzo	Matricola
Luca Neri	27/10/1980	Via Cavour 5, Milano	123456
Mario Rossi	4/7/1981	Via Roma 4, Pisa	123768
Paolo Bianchi	12/1/1983	Via Reno 2, Bologna	123987
:	:	:	:

Chiave Primaria

Come si indica una chiave primaria?

- Studenti {Nome, Indirizzo, Data di Nascita, Matricola}

A cosa serve una chiave primaria?

- La chiave primaria permette di identificare univocamente un record
- La chiave primaria permette di trovare un record in modo efficiente

Come si può identificare una chiave primaria?

- Dipende dal dominio applicativo

Chiave primaria (una definizione più precisa)

Un sottoinsieme proprio K di attributi tale che

1. non possono esistere due record distinti che hanno gli stessi valori nei campi relativi agli attributi in K (*unicità*)
2. non è possibile sottrarre un attributo a K senza che la condizione di unicità cessi di valere (*non ridondanza*)

Esempio

Verbali {Matricola, Esame, Data, Voto}

Relazioni (I)

Database dei clienti e dei conti intestati ai clienti
Un cliente può avere uno o più conti

NOME	CODICE FISCALE	INDIRIZZO	NUMERO DI CONTO	SALDO
...
M. Rossi	MRORSS57A09E625T	Via Roma, 4 Pisa	556	10M
...
M. Rossi	MRORSS57A09E625T	Via Roma, 4 Pisa	667	3.4M
⋮	⋮	⋮	⋮	⋮

Duplicazione dell'informazione: Svantaggi:

- Aumentano, inutilmente, le dimensioni della tabella
- Inserimenti e modifiche devono essere duplicati

Relazioni (II)

Una possibile soluzione...

NOME	...	INDIRIZZO	NC1	S1	NC2	S2
M. Rossi	...	Via Roma, 4 Pisa	556	10M	667	3,4M
⋮	⋮	⋮	⋮	⋮	⋮	⋮

...e se un cliente ha tre o più conti?

...la maggior parte dei clienti ha un solo conto!
(i campi vuoti occupano memoria)

Il problema è che si usa *una* sola tabella per descrivere *due* diverse entità: i clienti ed i conti

Relazioni (III)

CLIENTI

NOME	CODICE FISCALE	INDIRIZZO
⋮	⋮	⋮
Claudio Neri	CLDNRI68B12E625Z	Via Derna, 8 Pisa
Mario Rossi	MRORSS57A09E625T	Via Roma, 4 Pisa
⋮	⋮	⋮

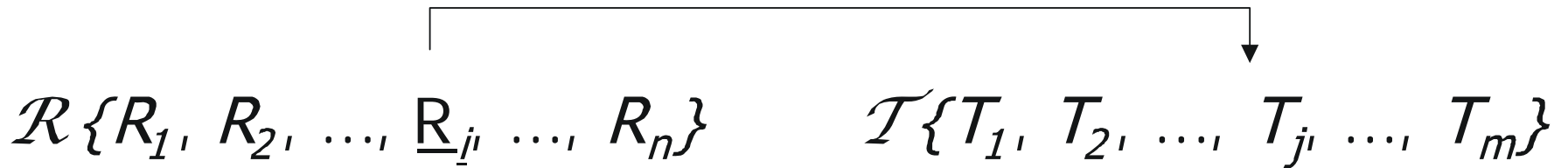
messi a comune

CONTI

CLIENTE	NUMERO DI CONTO	SALDO
⋮	⋮	⋮
MRORSS57A09E625T	556	10M
⋮	⋮	⋮
MRORSS57A09E625T	667	3.4M
⋮	⋮	⋮

Due tabelle il cui *tema* è unico

Chiave esterna



T_j è una *chiave esterna* della tabella \mathcal{T} proveniente dalla tabella \mathcal{R}

La tabella \mathcal{T} è la tabella *figlio*; la tabella \mathcal{R} è la tabella *padre*

La tabella figlio *dipende* dalla tabella padre per completare la definizione dei propri record

Tipi di Relazione

- **Relazione uno-a-molti**

- ad un record di una tabella possono corrispondere uno o più record della seconda

- **Relazione multi-a-molti**

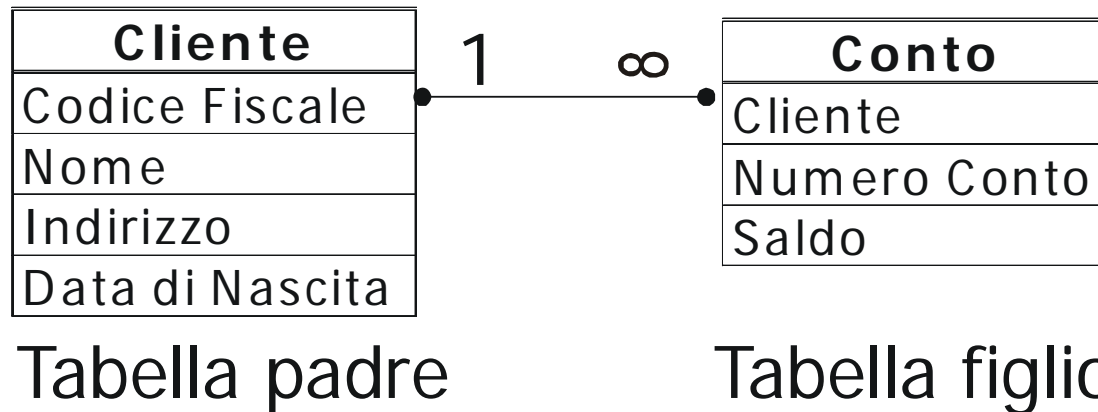
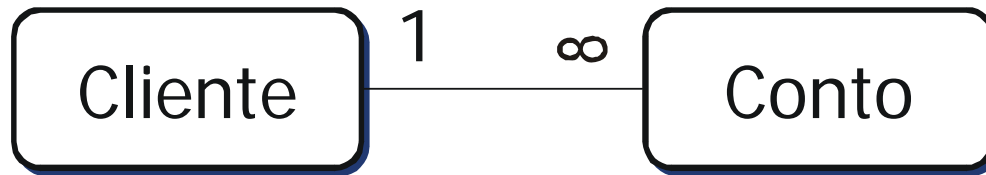
- ad un record di una tabella possono corrispondere uno o più record della seconda e viceversa

- **Relazione uno-a-uno**

- ad un record di una tabella può corrispondere al più un record della seconda e viceversa

Uno-a-Molti

Un cliente può avere uno o più conti; un conto può essere intestato ad un solo cliente



Multi-a-Multi

Un cliente può avere uno o più conti; un conto può essere intestato ad uno o più clienti

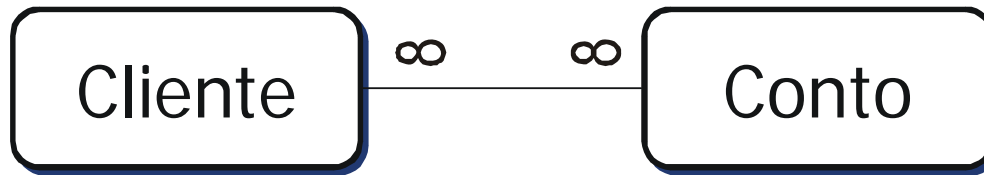
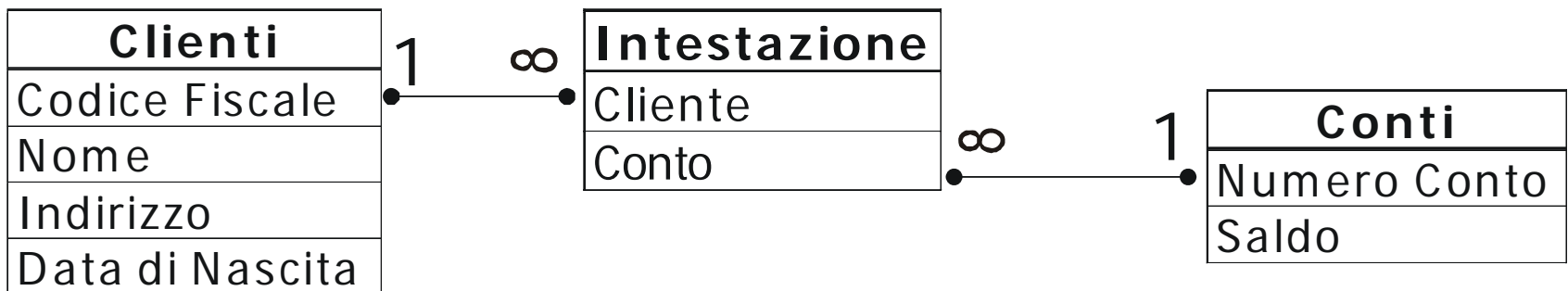
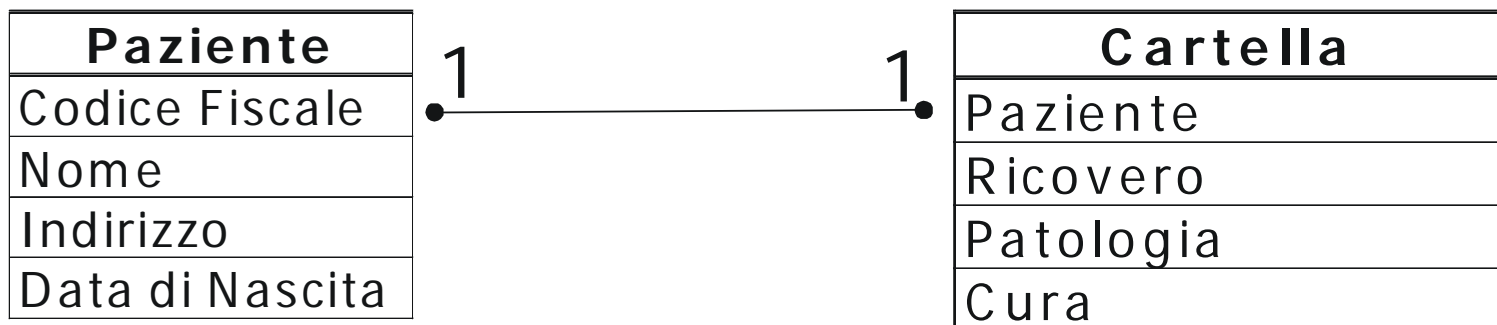


Tabella ponte



Uno-a-uno

Un paziente ha un'unica cartella medica; una cartella medica appartiene ad un solo paziente



Valori Nulli (I)

Si consideri la tabella

Cliente {Codice Fiscale, Nome, Indirizzo, Telefono}

Il cliente Mario Rossi non ha telefono...

...cosa si scrive nel campo Telefono?

Si prende un valore del dominio e gli si attribuisce un significato speciale

Questa soluzione non è "pulita" e non sempre possibile

Valori Nulli (II)

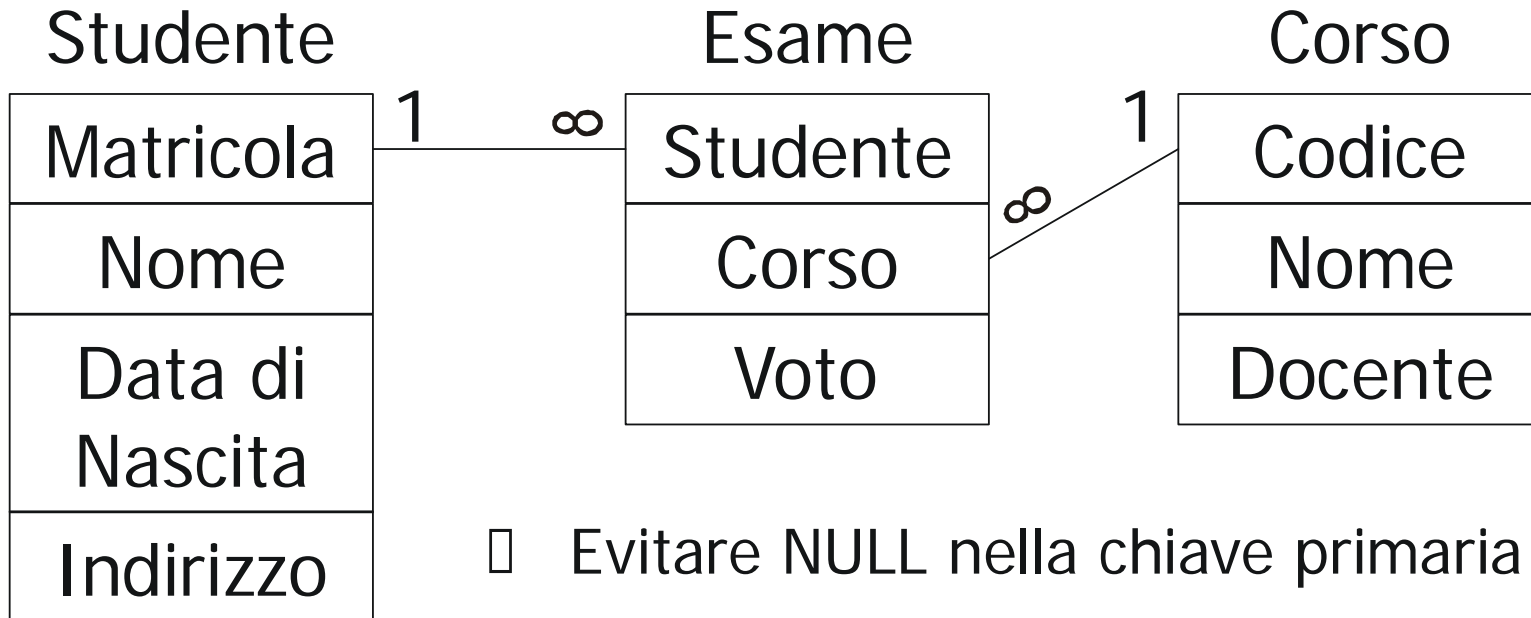
Una soluzione migliore consiste nel rendere esplicita l'assenza di valori:

Un campo può assumere o un valore del dominio o il valore NULL

Interpretazione del valore NULL:

- Informazione inesistente (l'utente non ha telefono)
- Informazione sconosciuta (il numero di telefono dell'utente non è noto)
- Entrambe

Valori Nulli (III)



- Evitare NULL nella chiave primaria
- Evitare NULL nella chiave esterna
- Evitare NULL multipli in un record

Sulla chiave primaria si vietano i valori NULL

Integrità dei Riferimenti

Mario Rossi cambia banca

CLIENTI

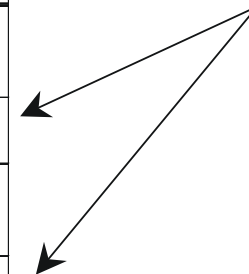
NOME	C ODICE FISCALE	INDIRIZZO
⋮	⋮	⋮
Claudio Neri	CLDNRI68B12E625Z	Via Derna, 8 Pisa
Mario Rossi	MRORSS57A09E625T	Via Roma, 4 Pisa
⋮	⋮	⋮

CONTI

CLIENTE	NUMERO DI CONTO	SALDO
⋮	⋮	⋮
MRORSS57A09E625T	556	10M
⋮	⋮	⋮
MRORSS57A09E625T	667	3.4M
⋮	⋮	⋮



Record orfani



Database

Regole di Integrità dei Riferimenti

Fanno sì che due tabelle in relazione tra loro rimangano *mutuamente* consistenti evitando il formarsi di record orfani

Regole

1. *Cancellazione in cascata*
2. *Modifica in cascata*
3. *Inserimenti correlati*