

«Di fronte alla giungla di documenti che si aggiungono minuto per minuto, la domanda cruciale è piuttosto la seguente: se lancio un'informazione in rete, qualcuno la noterà? Per esser letti bisogna essere visibili: una banale verità che vale tanto per gli scrittori quanto per gli scienziati.»

(Albert-László Barabási il «Link»)

Abilità Informatiche

Informazione: struttura, fruizione e distribuzione

Università degli Studi di Ferrara

Dipartimento di Studi Umanistici

Corsi di Laurea Magistrale in «Lingue e Letterature Straniere»

aa 2014-2015

Prof. Giorgio Poletti – giorgio.poletti@unife.it

I Data Base Gerarchici

Anni 60 – Struttura ad Albero

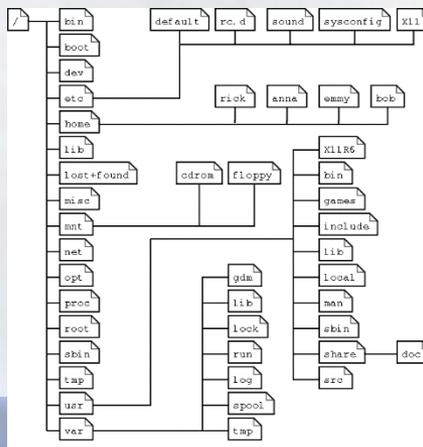
CARATTERISTICHE

- Gerarchia tra **entità**
- Una Radice e uno o più sottoalberi dipendenti
- Struttura **1:N (padre:figli)**

CRITICITÀ

- Albero come schema logico
- Scarsa indipendenza dello schema logico
- Gli schemi esterni devono esplicitare tutto il percorso

Esempio il file system degli HD



I Data Base Reticolari

Anni 70 – Struttura a Grafo

CARATTERISTICHE

- Record e Puntatori
- Struttura **M:N**

CRITICITÀ

- Aumento della quantità di dati e Ridondanza
- Aumento della possibilità di errore



I Data Base Relazionali

Anni 70 – Struttura a Tabelle

TABELLA = RELAZIONE

CARATTERISTICHE

- Tutti i dati sono relazioni e manipolati con l'**algebra relazionale**
- Struttura **M:N**

CRITICITÀ

- Problemi nella modellazione di relazioni complesse

*Teoria degli insiemi e logica del 1 ordine**

* *logica del 1 ordine* è linguaggio formale per la gestione **meccanica** degli enunciati che utilizza:

- *Connettivi logici* (AND, OR...)
- *Relazioni* (Contiene, è contenuto, è strettamente contenuto)
- *Quantificatori* (per ogni ... esiste)

Matricola	Data	Studente
10001	10/10/87	Paolo Rossi
10002	12/09/91	Enrico Verdi

Studente	Voto	Corso
10001	24	0006
10001	27	0002
10001	30	0004

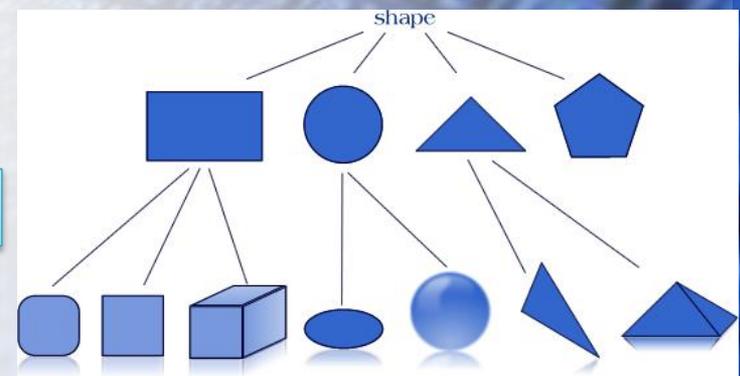
Codice	Titolo	Docente
0006	Analisi	Paolo Bianchi

I Data Base a Oggetti

Anni 80 – Struttura ad Oggetti

Paradigma **Object Oriented**: memorizzazione dei dati come «*oggetti riutilizzabili*». Oggetto come insieme di diversi tipi di **dati** (testo, audio, video) e **metodi** (operazioni eseguibili sui dati).

ADATTO A BASI DATI MULTIMEDIALI



I Data Base Management System (DBMS)

DBSM (Data Base Management System): sistema di controllo che permette la gestione di grandi masse di dati attraverso uno strumento centralizzato.

CARATTERISTICHE

un dato o una informazione siano rimasti integri sono rimasti inalterati nel loro contenuto, durante la loro trasmissione oppure la loro memorizzazione.

controllo della ridondanza dei dati e consistenza dei dati

accessi simultanei ai dati

integrità dei dati

Consistenza: sinonimo di coerenza, conformità, non contraddittorietà, affidabilità dei dati.

sicurezza delle informazioni

facile accesso ai dati per non esperti



I Data Base Management System (DBMS)

Astrazione

Livello Interno

Descrizione dei dati elementari e dei record di informazioni

Schema Fisico

Livello Logico

Descrizione della logica che governa il complesso delle informazioni

Schema Logico

Livello Esterno

Descrizione delle viste logiche sui dati che vengono rese disponibili per gli utenti e le diverse tipologie di utenti previste

Sottoschemi

I Data Base Management System (DBMS)

Schema di linguaggi e strutture per la gestione



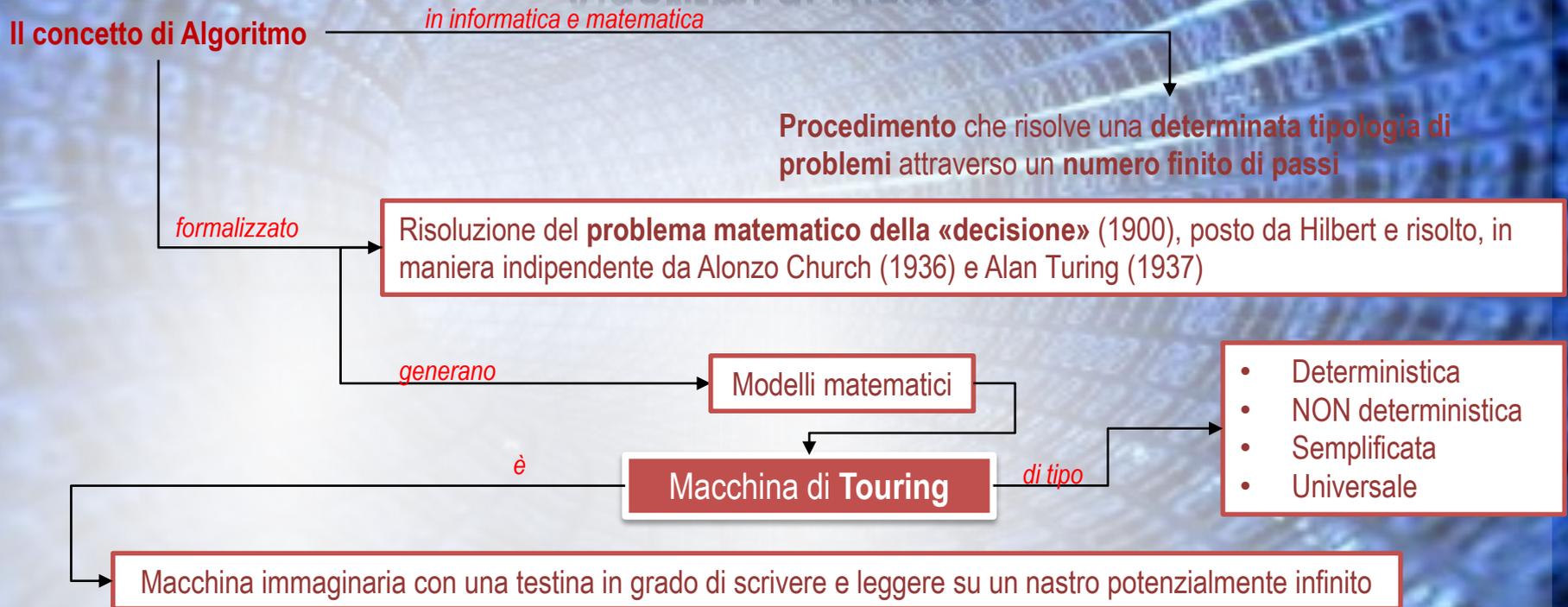
I Data Base Management System (DBMS)

Schema della struttura del software di gestione



I Data Base Management System (DBMS)

Algoritmi di Ricerca



I Data Base Management System (DBMS)

Algoritmi di Ricerca

Il concetto di Algoritmo

in informatica e matematica

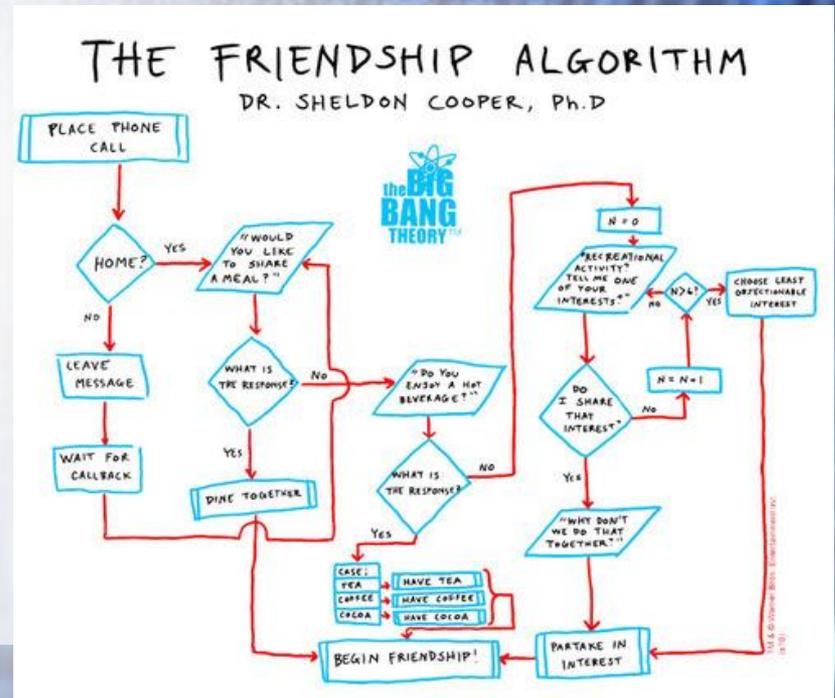
tipologia

Procedimento che risolve una **determinata tipologia di problemi** attraverso un **numero finito di passi**

consideriamo

Molteplici e rendono impossibile una catalogazione precisa e definitiva

- Algoritmi di ordinamento
- Algoritmi di ricerca
- Algoritmi Genetici (evolutivi)
- Algoritmi di visita degli alberi



I Data Base Management System (DBMS)

Algoritmi Genetici

- **Cromosoma:** una delle soluzioni ad un problema considerato. Generalmente è codificata con un vettore di bit o di caratteri.
- **Popolazione:** insieme di soluzioni relative al problema considerato.
- **Gene:** parte di un cromosoma. Generalmente consiste in una o più parti del vettore di bit o caratteri che codificano il cromosoma.
- **Fitness:** grado di valutazione associato ad una soluzione. La valutazione avviene in base ad una funzione appositamente progettata detta funzione di fitness.
- **Crossover:** generazione di una nuova soluzione mescolando delle soluzioni esistenti.
- **Mutazione:** alterazione casuale di una soluzione.

ESEMPIO DI ALGORITMO GENETICO (PROVEDE A FARE EVOLVERE DELLE SOLUZIONI)

1. Generazione casuale della prima popolazione di soluzioni (**CROMOSOMI**).
2. Applicazione della funzione di **FITNESS** alle soluzioni (cromosomi) appartenenti all'attuale popolazione.
3. Selezione delle soluzioni considerate migliori in base al risultato della funzione di fitness e della logica di selezione scelta.
4. Procedimento di **CROSSOVER** per generare delle soluzioni ibride a partire dalle soluzioni scelte al punto 2.
5. Creazione di una nuova **POPOLAZIONE** a partire dalle soluzioni identificate al punto 3.
6. Riesecuzione della procedura a partire dal punto 2 ed utilizzando la nuova **POPOLAZIONE** creata al punto 4.

L'iterazione dei passi presentati permette l'evoluzione verso una soluzione ottimizzata del problema considerato.