

INTRODUZIONE ALL'INFORMATICA



Prof. Donato Malerba
Dipartimento di Informatica
Università degli Studi Bari



UNIVERSITÀ
DEGLI STUDI DI BARI
ALDO MORO

I Seminari dell'Orientamento Consapevole
Bari, 18 gennaio 2016



DIPARTIMENTO
DI INFORMATICA

Informatica

L'informatica è la disciplina scientifica che studia tutto ciò che ha a che fare con il calcolabile.

In inglese si usa l'espressione "*computer science*", ponendo enfasi sullo strumento di calcolo.

Ma ...

L'informatica non riguarda i computer più di quanto l'astronomia riguardi i telescopi. (*Edsger Wybe Dijkstra*)



Informatica

In italiano (e in francese) si preferisce la parola "informatica", che nasce dalla contrazione di informazione e automatica, per indicare l'elaborazione automatica dell'informazione, che è appunto l'attività effettuata dal computer.

Ma che vuol dire che l'informatica è una disciplina scientifica, cioè una scienza?

Cosa è una disciplina?

Lat. **Disciplina(m)**, originariamente
`insegnamento, educazione', da **discere**
`imparare'

- Materia di studio – Ramo del sapere
(da Zanichelli, Vocabolario della lingua italiana, 1994)
- Quando un insieme di conoscenze diventa `disciplinato'?
- Quando queste sono generate mediante *paradigmi*

Cos'è un paradigma?

Paradigma: dal greco παραδειγμα 'modello, esempio', da παρα 'presso, accanto' (indica similarità) + δεικνυvai 'mostrare, indicare'

- Sebbene il termine paradigma sia frequentemente assimilato al concetto di modello, esso è caratterizzato dal fatto che può essere utilizzato per *produrre* qualcosa.
- Esempio: paradigma dei verbi.

Cos'è un paradigma?

- Nella scienza un paradigma è un modello utilizzato per produrre nuova conoscenza.
- Esso si connota come una teoria aperta, ampia, verificabile, senza precedenti che getta le fondamenta di una disciplina.

Thomas Kuhn (filosofo, 1922-1996)

I paradigmi aiutano le comunità scientifiche a delimitare la loro disciplina in quanto aiutano lo scienziato a

- creare percorsi di ricerca
- formulare domande
- selezionare metodi con i quali esaminare le domande
- definire aree di rilevanza
- [stabilire/creare significati?]



Thomas Kuhn (filosofo, 1922-1996)

- “In assenza di un paradigma o di qualche candidato a paradigma, tutti i fatti che potrebbero riferirsi allo sviluppo di una data scienza appariranno molto probabilmente ugualmente rilevanti.”
- “Una **rivoluzione** (scientifica) è un cambio di paradigma”.



Paradigmi per la costruzione della conoscenza

- Teoria
- Astrazione / Sperimentazione

Paradigma della teoria

Ha le sue radici nella **matematica** ed è costituito da quattro passi che vengono seguiti nello sviluppo di una teoria valida e coerente:

1. Caratterizzare gli oggetti dello studio (definizione)
2. Ipotizzare possibili relazioni fra loro (teorema)
3. Determinare se le relazioni sono vere (dimostrazioni)
4. Interpretare i risultati

Paradigma della teoria

Esempio:

1. Definizione di relazione di congruenza, esponente, modulo e numero primo
2. Piccolo teorema di Fermat (enunciato da Fermat, ma mai dimostrato): se p è un numero primo ed a un qualunque intero non divisibile per p , allora $a^{p-1} - 1$ è divisibile per p .

$$a^p \equiv a \pmod{p}$$

3. Dimostrazione del teorema (es., quella di Leibniz)
4. Il teorema di Fermat, dá una condizione necessaria per verificare che un numero sia primo (test di primalità di Fermat).

Paradigma dell'astrazione (o sperimentazione)

Ha le sue radici nel **metodo scientifico sperimentale** (tipico della fisica) ed è costituito da quattro passi che vengono seguiti nell'investigazione di un fenomeno:

1. Formare un'ipotesi
2. Costruire un modello e fare una previsione
3. Disegnare un esperimento e raccogliere i dati
4. Analizzare i risultati

Paradigma dell'astrazione (o sperimentazione)

Esempio:

- 1. Ipotesi:** Bere vino influisce sul sistema cardiocircolatorio
- 2. Modello:** Il rischio cardiaco è inversamente lineare all'assunzione di vino. Predizione: meno casi di infarti per chi beve di più, a parità di condizioni (età, ambientali, etc.).

Paradigma dell'astrazione (o sperimentazione)

Esempio:

- 3. Progetto dell'esperimento.** Assumendo che le cavie abbiano una risposta simile a quella dell'uomo, si decide di somministrare giornalmente del vino a 20 cavie scelte casualmente fra quelle della stessa età e di osservare a distanza di un anno le condizioni del sistema circolatorio mediante un ECG. Le cavie sono divise in 4 gruppi, caratterizzati dalla quantità di vino somministrata. Le cavie vivranno nello stesso ambiente e saranno alimentate allo stesso modo (controllo di fattori esterni). Si conduce l'esperimento e si raccolgono i dati.

Paradigma dell'astrazione (o sperimentazione)

Esempio:

- 4. Analisi dei risultati** con metodi statistici. Si osservi che il risultato non è deterministico perché il fattore genetico non è controllato/controllabile e potrebbe essere la fonte di variabilità nella risposta.

Curiosità

- Il paradigma sperimentale è stato il primo a essere applicato (già da alcune migliaia di anni).
- Solo alcune centinaia di anni fa si è cominciato ad applicare il paradigma della teoria per produrre nuova conoscenza (leggi di Keplero, leggi di Newton, equazioni di Maxwell).

Metodiche della progettazione

Gli oggetti di studio dell'informatica non sono né naturali né ideali: essi vanno progettati.

Questo comporta la necessità di applicare delle metodiche tipiche dell'ingegneria. Anche qui abbiamo quattro passi che vengono seguiti nella costruzione di un sistema atto a svolgere un dato compito:

1. Formulare i requisiti di stato
2. Specificare lo stato
3. Progettare e implementare il sistema
4. Test del sistema

Informatica come disciplina

Le conoscenze informatiche sono in verità generate da un mix dei due paradigmi (teoria e sperimentazione).

Gli oggetti di studio (programmi, sistemi informatici, etc.) sono costruiti applicando metodiche ingegneristiche.

Informatica come disciplina

Il paradigma della teoria ha contribuito alla maturazione delle conoscenze su computabilità, complessità computazionale, algoritmi e strutture dati, automi, linguaggi formali.

Il paradigma dell'astrazione ha contribuito allo studio di metriche per programmi e sistemi, alle simulazioni di sistemi e di processi fisici, all'analisi delle prestazioni di sistemi, al test di protocolli, al confronto delle architetture, all'apprendimento automatico.

Le metodiche della progettazione hanno contribuito allo sviluppo di ambienti di sviluppo, simulatori, sistemi CAD, CAM, VLSI, database, sistemi operativi.

Informatica come disciplina

Nei corsi di programmazione e di algoritmi e strutture dati si insegna a progettare delle soluzioni a partire da delle specifiche, a implementarle e a testarle. →

metodiche della progettazione

Nel corso di architettura e di sistemi operativi si insegna a costruire il modello di un calcolatore a vari livelli di astrazione, a valutare l'efficienza di una istruzione o di un insieme di istruzioni, o di un sistema di gestione della memoria. →

**paradigma dell'astrazione +
metodiche della progettazione**

Nel corso di basi di dati si insegna a progettare il modello concettuale e logico di una base di dati. → **metodiche della progettazione**

Informatica come disciplina

Nei corso di ingegneria del software si insegna a il software → **metodiche della progettazione**
ma anche a raccogliere dati sperimentali per analizzare delle proprietà del software → **paradigma della sperimentazione**

Nel corso di linguaggi di programmazione si studiano i modelli (automi e grammatiche) in grado di tradurre, riconoscere e generare dei linguaggi → **paradigma della teoria**

Nel corso di fondamenti dell'informatica si studiano i modelli di calcolo, i limiti della calcolabilità, la complessità computazionale di algoritmi e di problemi → **paradigma della teoria**

Paradigmi non tradizionali

L'informatica non ha solo 'preso in prestito' dei paradigmi e delle metodiche da altre discipline scientifiche e ingegneristiche.

L'informatica ha contribuito allo sviluppo di nuovi paradigmi per l'indagine scientifica!

Per questo essa è ormai 'pervasiva' e 'trasversale'.

Paradigmi non tradizionali

- Per molti problemi, i modelli teorici sono diventati molto complicati per essere risolti analiticamente, per cui ci si è affidati alla simulazione.
- La simulazione mediante calcolatore (**computational simulation**) ha rappresentato un nuovo paradigma per la produzione di conoscenza negli ultimi 50 anni (Ken Wilson, premio Nobel della fisica, 1982).
- Si parla di **terzo paradigma** della scienza a supporto della teoria e della sperimentazione.

Computational simulation

I passi fondamentali in cui si articola l'applicazione di questo paradigma sono:

1. Astrai le componenti di un sistema complesso e le interazioni fra le componenti e parametrizzale
2. Descrivi le componenti e le interazioni mediante un programma
3. Genera dei dati al variare dei parametri
4. Osserva le macro proprietà del sistema e interpretale

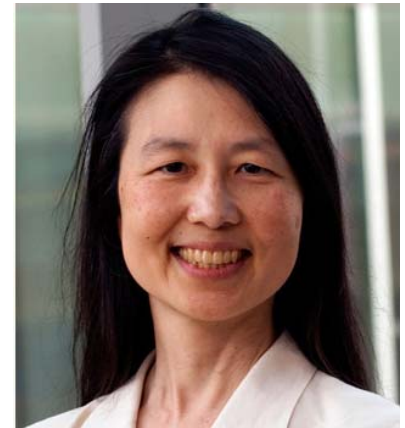
La nuova conoscenza è data proprio dalle macro proprietà osservate (e.g., evoluzione dinamica).

Computational simulation

- Le simulazioni si utilizzano per i crash test delle automobili, per studiare l'aerodinamicità di una struttura, e in generale per studiare le proprietà di un sistema complesso.
- Questo terzo paradigma è evidente in tutte le *computational sciences* emerse negli ultimi anni.
- Esempio: la *computational biology*, che simula il comportamento di sistemi biologici e delle vie metaboliche (*metabolic pathways*) o il comportamento di una cellula o il modo in cui viene costruita una proteina.

Computational simulation

Per dirla con Jeannette Wing
(vicepresidente Microsoft Research),
dietro l'emergere di queste scienze
computazionali c'è l'influenza del
“**Computational Thinking**”.



<http://www.cs.cmu.edu/afs/cs/usr/wing/www/publications/Wing06.pdf>

<http://www.cs.cmu.edu/~CompThink/>

Un quarto paradigma?

Recentemente Gordon Bell, che come a Ken Wilson ha sostenuto il riconoscimento del terzo paradigma, ha indicato un nuovo paradigma alla base della scienza moderna:

la **data intensive scientific discovery**.

<http://research.microsoft.com/en-us/collaboration/fourthparadigm/>



Un quarto paradigma?

- In molte discipline si raccolgono grandi collezioni di dati che non verranno mai analizzati. La raccolta di dati non è più finalizzata a un esperimento, come nel caso del paradigma della sperimentazione o nella simulazione. In molti progetti, la raccolta dei dati diventa un fine e non un mezzo per produrre conoscenza.
- Purtroppo molti dei dati raccolti non serviranno mai a produrre nuova conoscenza.
- Se negli ultimi anni sono stati fatti dei progressi nella costruzione di strumenti di simulazione, siamo ancora agli inizi nella produzione di efficaci strumenti di analisi.

Un quarto paradigma?

- Il quarto paradigma promette la produzione di nuove ipotesi a partire da queste vaste raccolte di dati, ipotesi che potranno essere poi verificate mediante strumenti analitici, sperimentali e di simulazione.
- Quindi, dopo l'osservazione empirica, volta alla descrizione dei fenomeni naturali; la riflessione teorica, che mira a generalizzare i risultati dell'osservazione e costruire modelli; e negli ultimi decenni, l'approccio computazionale, che costruisce la simulazione dei fenomeni complessi, si sta recentemente affermando l'esplorazione e manipolazione di grandi quantità di dati, come un nuovo paradigma.

Un quarto paradigma?

I passi in cui si articola l'applicazione del quarto paradigma sono:

- 1. Capture:** i dati sono "catturati" da sensori o strumenti e archiviati
- 2. Curate:** i dati sono gestiti nel tempo in modo da garantire la loro effettiva usabilità nel tempo. Per dataset dinamici questo è molto importante (e.g., l'adozione di un identificatore universale per un gene richiede una reindicizzazione dei geni memorizzati un database, altrimenti i dati non sarebbero incrociabili con altri).
- 3. Analyze:** i dati sono analizzati con metodi di data mining e knowledge discovery
- 4. Publish:** i risultati delle analisi sono interpretati e resi pubblici (è importante ricorrere a tecniche di visualizzazione)

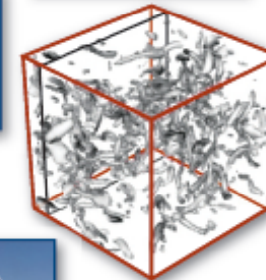
Riepilogando ...

Science Paradigms

- Thousand years ago:
science was **empirical**
describing natural phenomena
- Last few hundred years:
theoretical branch
using models, generalizations
- Last few decades:
a **computational** branch
simulating complex phenomena
- Today: **data exploration** (eScience)
unify theory, experiment, and simulation
 - Data captured by instruments
or generated by simulator
 - Processed by software
 - Information/knowledge stored in computer
 - Scientist analyzes database/files
using data management and statistics



$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - K\frac{c^2}{a^2}$$



Tutti i paradigmi sono utili ...

... anche se ci sono posizioni provocatorie come quella di C. Anderson che in

"The End of Theory: The Data Deluge Makes the Scientific Method Obsolete, Wired, June 2008

predice la fine del modello di scienza dominante:

"La correlazione sostituisce la causalità e la scienza può progredire anche senza modelli coerenti, senza teorie unificate o addirittura senza alcun modello di spiegazione. Non ha alcun senso aggrapparsi alle nostre vecchie pratiche. È il momento di domandarsi: cosa può imparare la scienza da Google?"

Non c'è più scienza senza l'informatica?

