



UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Do we need data structures? The top data structure you should know for next generation computing

1920-87R-03

Obiettivi

Il corso presente alcuni approcci recenti al disegno di strutture dati per trattare sfide computazionali moderne relative alla computazione su grandi moli di dati. Obiettivo del corso è l'acquisizione di alcune tecniche e di alcuni concetti che consentono di affrontare sfide emergenti nell'ambito Computer Science.

In particolare il focus è su:

- le applicazioni moderne che richiedono queste strutture dati (es. Web, Bioinformatica, Data Mining)
- *Algoritmi su testi e data streaming*

Contenuti sintetici

Alcune lezioni sono tenute da esperti dell'ambito

In particolare secondo il seguente calendario:

Travis Gagie terrà un seminario il 9 Marzo -

Lezioni principali - terza e quarta settimana di Marzo -

Introduzione

- Nuove strutture dati e loro utilizzo in diverse applicazioni

- cosa si intende per indicizzazione di big data?

quali sono le principali query sull'indice?

Che cosa è il data streaming?

- Introduzione ai Bloom Filters e la loro applicazione per gestire big data in poco spazio

PARTE 2 (data prevista il 25 Marzo)

Tecniche di hashing con applicazioni al WEB

(Web crawling and indexing, ...)

(parte tenuta da Paolo Boldi <http://boldi.di.unimi.it/>)

MPHF (minimum perfect hashing function)

- Prima parte: (M)PHF

? Il problema generale (principale applicazione: la costruzione del grafo)

? funzioni di hash

? L'algoritmo MWHC per array sparsi

? MWHC per costruire (M)PHF

- Seconda parte: MMPHF

? Introduzione

? Primo tipo di costruzione: LCP

? Secondo tipo di costruzione : Z-fast tries

PARTE 3 (dal 24 Marzo)

Strutture dati per il pattern matching

Suffix-trees e suffix-arrays

(parte tenuta da Gregory Kucherov, docente del collegio di dottorato

<https://www-igm.univ-mlv.fr/~koutcher>)

Programma esteso

Prerequisiti

Modalità didattica

Materiale didattico

Slides distribuite a lezione e usate per le lezioni + eventuali articoli

Periodo di erogazione dell'insegnamento

Modalità di verifica del profitto e valutazione

L'esame consistente nel fare una breve presentazione specifica su un argomento di interesse che è stato svolto durante il corso

Orario di ricevimento
