DM803 – Advanced Data Structures

(Avancerede Datastrukturer)

"the pizza meeting"

Kim Skak Larsen

IMADA
University of Southern Denmark

kslarsen@imada.sdu.dk

November 18, 2021

Example Considerations

You know that operations on red-black trees are $O(\log n)$. However,

- $O(\log n)$ comes from searching – what about just rebalancing?
- how many of the rebalancing operations change pointers?
- are any changes possible to obtain $O(1)$ rebalancing?
- its precise height guarantee is $2 \log n$ – can any BST do better?

You know the heap implementation of a priority queue – what if we want to meld two priority queues (efficiently)?

Can we keep dynamic structures balanced *without* any balance information?

Are there realistic computational models where $\Omega(n \log n)$ is *not* a lower bound on sorting?

Course Contents

- priority queues
- height- and weight-balanced trees
- multi-way trees
- randomized search structures
- disjoint sets with variations
- hashing methods
- techniques such as
  - amortized analysis
  - persistency
  - dynamization & global rebuilding
  - word RAM manipulations – beating $\Omega(n \log n)$ for sorting

More Concretely

- leftists heaps, skew heaps, and Fibonacci heaps
- skip lists
- scapegoat trees
- cuckoo hashing
- analysis of disjoint sets
- disjoint sets with backtracking
- making data structures partially persistent
- van Emde Boas trees, X- and Y-fast tries
- splay trees
- AVL trees
- treaps
- (a, b)-trees
- level ancestors
- relaxed balance

Example Applications

Data structures are important parts of:

- Efficient algorithms for fundamental problems in CS
- Database Management Systems
- Geographic Information Systems (GIS)
- Compilers/interpreters for various programming languages
- Robot Motion Planning
- Computer Aided Design
- Games
- ...

(But we are not looking at applications in this course.)

Prerequisites

Topics from algorithms and data structures (DM507) and discrete math (DM547/DM549/MM537/MM540) as well as programming experience. Some theory "maturity" from DM551 and/or DM553 would be great!

Specific Data Structures

- Search Trees (red-black trees)
- Priority Queues (binary heap)
- Disjoint Sets (Galler-Fischer representation)

General Techniques

- Asymptotic Notation
- Time and Space Analysis
- Induction
- Probability Theory

Formal Details

- 10 ECTS
- throughout the spring semester
- 2 hours of lectures and 2 hours of exercises per week
- articles and excerpts from textbooks
- exam project throughout (implementation of data structures)
- oral exam with preparation
- combined evaluation of the two parts (7 point scale)

Disclaimer

- There will be some theorems and proofs in almost every lecture
- – in particular, careful analysis of running time
- We will not look much at applications
- Chalk & blackboard lectures
- Course language is English, if necessary
- There will not be a TA helping you with debugging when you program
- Some maturity from DM553 or math courses would not hurt

In Doubt?

I would have liked to be present at the pizza meeting, but that was unfortunately not possible at the given time.

Please send me an email or – even better – drop by my office if you have questions.

I'm almost always here (from around 7:30). I can of course be away from my office at times, but I should normally be easy to find. And after my first cup of coffee, I don't bite...