Master Thesis Preparation

Algorithms

Gerth Stølting Brodal
Overview

Algorithms group at DAIMI
- Who?
- Where?
- Courses
- Research

Master thesis in Algorithms
- Types of thesis
- Recent thesis topics
Algorithms Group – Who?

Faculty
Lars Arge
Gerth Stølting Brodal
Gudmund Skovbjerg Frandsen
Peter Bro Miltersen
Christian Nørgård Storm Pedersen
(Erik Meineche Schmidt)
(Sven Skyum)

Researchers
Thomas Mailund
Herman Haverkort

Ph.d. students
7

Master students
21
Algorithms Group – Where?

Algorithms (Turing 2)
Arge, Brodal, Frandsen, Miltersen, Haverkort

BioInformatics
(Building 090)
Pedersen, Mailund
Algorithms Group – Courses

Introductory
• Programming 2 - Frandsen
• Algorithms and data structures - Brodal, Schmidt
• Machine architecture/Operating systems - Pedersen

Advanced
• Optimization/Combinatorial search - Miltersen
• Computational geometry - Arge, Brodal
• I/O algorithms - Arge, Brodal
• Dynamic algorithms - Frandsen
• Randomized algorithms - Frandsen
• String algorithms - Pedersen
• Algorithms in bioinformatics - Pedersen
• Complexity theory - Miltersen
• Compression - Miltersen
• Strategic game playing - Miltersen
Algorithms Group – Research

- I/O algorithms
- Computational geometry
- Data structures
- String algorithms
- Complexity theory
- Compression
- Optimization
- Algebraic algorithms
- BioInformatics
- Graph algorithms
- Dynamic algorithms
- Randomized algorithms

Subset of research interests
Solid lines = major interest
Algorithms Group – Research

- Theoretical computer science
- Tool development
  - BioInformatics, I/O algorithms
- Algorithm engeneering
  - primarily in relation to thesis work
- Algorithms and complexity research seminar
  - www.daimi.au.dk/~gerth/alcom-seminar/
### Results

<table>
<thead>
<tr>
<th>Problem</th>
<th>Best cache-oblivious result</th>
<th>Best cache-aware result</th>
</tr>
</thead>
<tbody>
<tr>
<td>List ranking</td>
<td>$O(\text{Sort}(V))$</td>
<td>$O(\text{Sort}(V))$</td>
</tr>
<tr>
<td>Euler Tour</td>
<td>$O(\text{Sort}(V))$</td>
<td>$O(\text{Sort}(V))$</td>
</tr>
<tr>
<td>Spanning tree/MST</td>
<td>$O(\text{Sort}(E) \cdot \log \log V)$</td>
<td>$O(\text{Sort}(E) \cdot \log \log (VB/E))$</td>
</tr>
<tr>
<td></td>
<td>$O(\text{Sort}(E))$ (randomized)</td>
<td>$O(\text{Sort}(E))$ (randomized)</td>
</tr>
<tr>
<td>Undirected BFS</td>
<td>$O(V + \text{Sort}(E))$</td>
<td>New $O(\text{ST}(E) + \text{Sort}(E) + \sqrt{VE/B})$</td>
</tr>
<tr>
<td></td>
<td>$O(\text{ST}(E) + \text{Sort}(E))$</td>
<td>New $O(\text{ST}(E) + \text{Sort}(E) + \sqrt{VE/B})$</td>
</tr>
<tr>
<td></td>
<td>$+ \frac{B}{V} \cdot \log V + \sqrt{VE/B}$</td>
<td>$+ \sqrt{VE/B} \cdot \sqrt{VB/E}^e$</td>
</tr>
<tr>
<td>Directed BFS &amp; DFS</td>
<td>$O((V + E/B) \cdot \log V + \text{Sort}(E))$</td>
<td>$O((V + E/B) \cdot \log V + \text{Sort}(E))$</td>
</tr>
<tr>
<td>Undirected SSSP</td>
<td>$O(V + (E/B) \cdot \log (E/B))$</td>
<td>New $O(V + (E/B) \cdot \log (E/B))$</td>
</tr>
</tbody>
</table>

Table 1. I/O-bounds for some fundamental graph problems.
Algorithm Research — another typical result

Comparisons by Quicksort
Element swaps
Running time

Figure 5: The number of comparisons, the number of element swaps, the number of branch mispredictions, and the running time of randomized Quicksort on F4, for \( n = 2 \times 10^5 \). The x-axis shows log10(n).
Master Thesis in Algorithms

Types of thesis

- Survey of a research area
- Implement a technical paper
  ...fill in the missing details
  ...perform experiments
- Explain all (missing) details in a technical paper
  ...how 8 pages becomes +100 pages
- Experimental comparison of several algorithms
- The good idea: Describe a new algorithm
Master Thesis in Algorithms

Thesis work

• Large fraction of time spend on trying to understand technical complicated constructions

• Implementations are often an ”existence proof” – most algorithm authors do not implement their algorithms (did they ever think about the missing details?)

• Hard to convince friends that it took you a year to understand an 8 page paper...
Hidden work...

Compact Oracles for Reachability and Approximate Distances in Planar Digraphs

Mikkel Thorup
AT&T Labs - Research, Shannon Laboratory
180 Park Avenue, Florham Park, NJ 07932, USA
mthorup@research.att.com

! Warning!
Nontrivial construction ahead of you

! Warning!
Need to understand another paper first

Proof: The proof is contained in [13], but somewhat hidden in other details because they need to ensure that the paths are of \( O(\sqrt{n}) \) length. The existence of \( v \) and \( w \) is what is actually proved in the proof of Lemma 2 in [13]. They find \((v, w)\) as an edge in an arbitrary triangulation of \( H \). No side of the fundamental cycle of \((v, w)\) in \( T \) contains more than \( 2/3 \) of \( H \). The vertices \( v \) and \( w \) are found in linear time in steps 1, 8, and 9 of the partitioning algorithm in §3 in [13].
## Recent Algorithms Master Theses

<table>
<thead>
<tr>
<th>Title</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refined Buneman Trees</td>
<td>Pedersen</td>
</tr>
<tr>
<td>Integer Sorting</td>
<td>Fagerberg</td>
</tr>
<tr>
<td>Trade-offs for Internal and External Memory Dictionaries</td>
<td>Fagerberg</td>
</tr>
<tr>
<td>A Survey of Density Keeping Algorithms</td>
<td>Fagerberg</td>
</tr>
<tr>
<td>Shortest Paths in Directed Graphs</td>
<td>Fagerberg</td>
</tr>
<tr>
<td>Approksimative afstande i planare grafer</td>
<td>Brodal</td>
</tr>
<tr>
<td>Vedligeholdelse af sammenhængskomponenter i dynamiske grafer</td>
<td>Frandsen</td>
</tr>
<tr>
<td>Maksimale par og suffikstræer</td>
<td>Pedersen</td>
</tr>
<tr>
<td>Skjulte Markov modeller og genidentifikation</td>
<td>Pedersen</td>
</tr>
<tr>
<td>Towards practical deterministic extractors</td>
<td>Miltersen</td>
</tr>
<tr>
<td>Engineering cache-oblivious sorting algorithms</td>
<td>Fagerberg/Brodal</td>
</tr>
<tr>
<td>Analyse og håndtering af geneekspressionsdata</td>
<td>Pedersen</td>
</tr>
<tr>
<td>Dynamisk Pattern Matching</td>
<td>Frandsen</td>
</tr>
<tr>
<td>Redigeringsafstande imellem niveau-strenge</td>
<td>Frandsen</td>
</tr>
<tr>
<td>Automated Layout of Classified Ads</td>
<td>Brodal</td>
</tr>
</tbody>
</table>