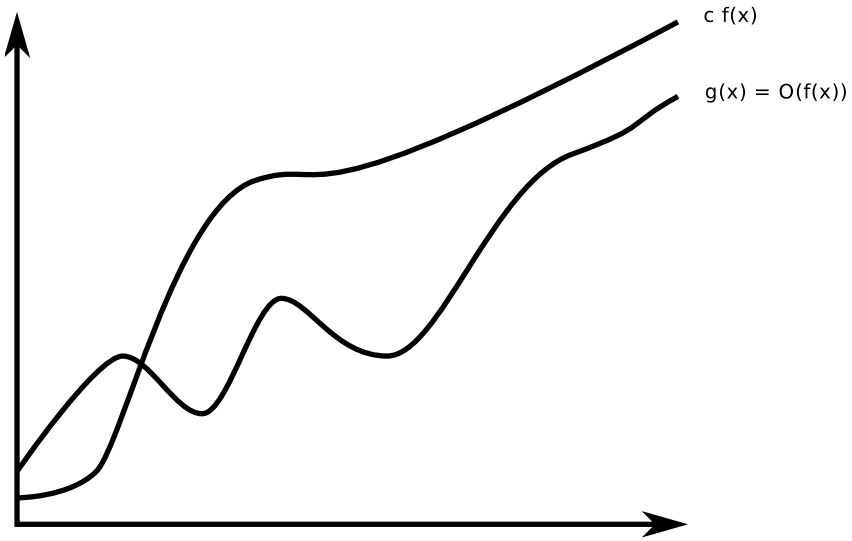
TDDD56: Multicore and GPU programming
Theory definitions

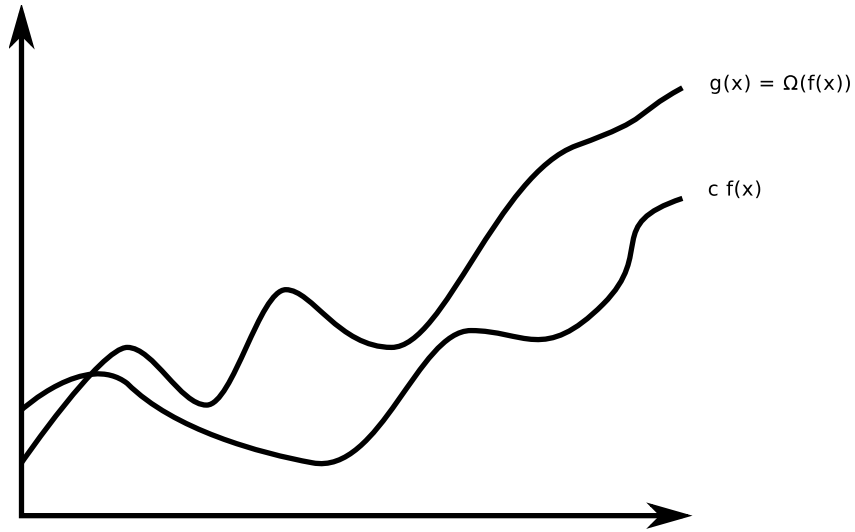
Nicolas Melot nicolas.melot (at) liu.se

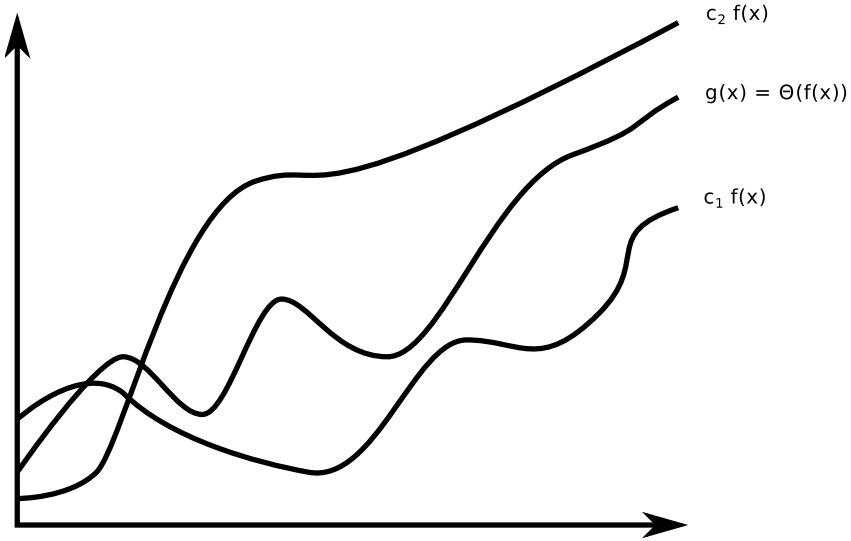
Linköping university (Sweden)

September 2, 2013

Definitions: Big-O notation



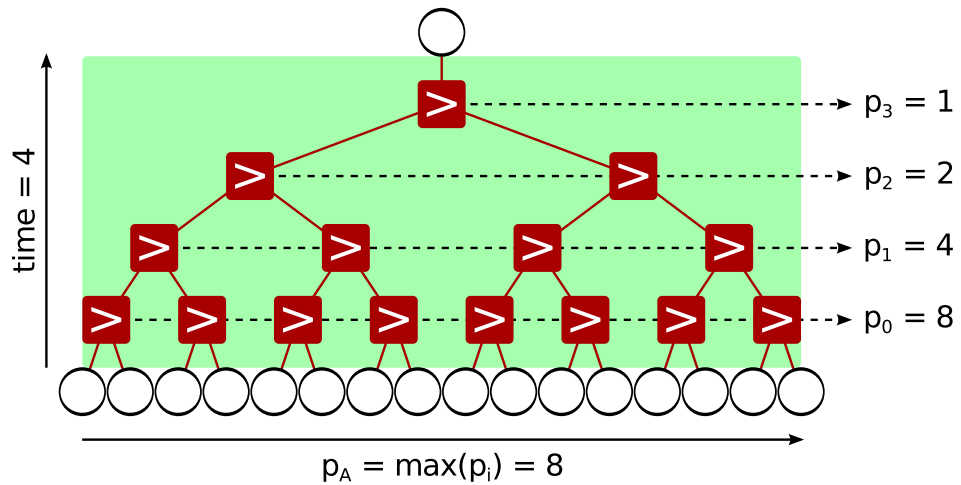




Definitions: metrics

- Let A be a parallel algorithm which best sequential equivalent runs in $t_{seq}(n)$ time units.
- Let $p_A = max(p_{A,i})$
- Work $w_A(p_A, n)$ = number of (non-idle) instructions for n-long input using p_A processors
  - Red surface, each square is unit
  - Red links included, depending on execution model
- Cost $c_A(p_A, n) = p_A \cdot t_A(P_A, n)$
  - Green surface

- Work optimality:
  $w_A(P_A, n) = O(t_{A,seq}(n))$
- Cost effectiveness:
  $c_A(p_A, n) \leq w_A(P_A, n)$
- Cost optimality:
  $c_A(p_A, n)) = O(t_{A,seq}(n))$



Definitions: parallel machines

- RAM: Random Access Machine
- PRAM: Parallel Random Access Machine
  - EREW PRAM: Exclusive Read, Exclusive Write PRAM
  - CREW PRAM: Concurrent Read, Exclusive Write PRAM
  - CRCW PRAM: Concurrent Read, Concurrent Write PRAM
    - ★ Weak: all parallel writes must be writing 0, or incorrect algorithm
    - ★ Common: all parallel writes must be writing the same value, or incorrect algorithm
    - ★ Abritrary: one rule defines which processor can write
    - ★ Priority: priorities are statically set to core; the higher priority core can write
    - ★ Combining: apply reductions to concurrent writes: add, min/max, bitwise OR, etc