Mastermath midterm examination
Parallel Algorithms

Teacher: Rob H. Bisseling, Utrecht University

October 26, 2022

Each of the four questions is worth 10 points. Total time 120 minutes. Closed-book exam. Motivate you answers!

1. (a) [5 pt] What is a *mixed superstep* in a BSP algorithm?

   (b) [5 pt] Explain the BSP cost of a mixed superstep. Give the meaning of the variables you use in the explanation.

2. Let $\mathbf{x}$ be a given vector of length n, which is distributed by the block distribution over p processors, with $n \bmod p = 0$. We want to compute the vector $\mathbf{y}$, also of length n, defined by

$$y_i = \max\{x_j \ : \ 0 \le j \le i\} \qquad \text{for } 0 \le i < n.$$

   (a) [5 pt] Give an efficient BSP algorithm for processor $P(s)$ (in the notation we learned) for the computation of the vector $\mathbf{y}$. The output vector $\mathbf{y}$ must be obtained in the block distribution.

   (b) [5 pt] Analyse the BSP cost of your algorithm. Here, assume that a comparison costs one flop, but an assignment is for free.

3. Let A be an $n \times n$ matrix, which is distributed by a square Cartesian distribution with $p = M^2$ processors. Assume that n and M are powers of 2 with $n > M$. Furthermore, assume we write A in block form as

$$A = \begin{bmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{bmatrix},$$

   where $A_{00}, A_{01}, A_{10}, A_{11}$ are matrices of size $n/2 \times n/2$.

   (a) [3 pt] We want to compute the matrix

$$B = \begin{bmatrix} A_{00} + A_{11} & A_{10} - A_{01} \\ A_{10} + A_{01} & A_{00} - A_{11} \end{bmatrix}.$$

   Choose an efficient distribution for A and B, either the square block distribution or the square cyclic distribution, with $\mathrm{distr}(A) = \mathrm{distr}(B)$. Be sure to motivate your choice.

(b) [4 pt] Describe a parallel algorithm for the computation of B with your chosen distribution. Use the notation we have learned to express algorithms, and use a_{ij} and b_{ij} to denote the elements of the matrices A and B, respectively.

(c) [3 pt] Analyse the BSP cost of your algorithm.

4. *Cholesky decomposition* is the equivalent of LU decomposition for symmetric matrices A; it computes L with $A = LL^{\mathrm{T}}$. If A is symmetric positive definite (SPD), the algorithm does not need pivoting. For this situation, the sequential algorithm is given by Algorithm 1. In the algorithm, A being SPD guarantees that the square root is always taken from a positive value a_{kk}.

Algorithm 1 Sequential Cholesky decomposition.

Input: A : $n \times n$ matrix, $A = \mathrm{Lower}(A^{(0)})$, i.e. the lower triangular part.
Output: A : $n \times n$ matrix, $A = L$, with
L : $n \times n$ lower triangular matrix, such that $LL^{\mathrm{T}} = A^{(0)}$.

```
for k := 0 to n − 1 do
    a_kk := √a_kk
    for i := k + 1 to n − 1 do
        a_ik := a_ik / a_kk;

    for j := k + 1 to n − 1 do
        for i := j to n − 1 do
            a_ij := a_ij − a_ik a_jk;
```

(a) [3 pt] Formulate (in the notation we have learned) the computation superstep of a BSP algorithm for processor $P(s,t)$ that parallelises the matrix update (the last three lines) of Algorithm 1. Assume a square $M \times M$ cyclic distribution for the matrix, with $p = M^2$.

(b) [2 pt] Compare the BSP cost of the matrix update in stage k to that of the LU decomposition. You may give the result as an approximation.

(c) [3 pt] Describe (in words) an efficient communication superstep preceding the matrix update that obtains the necessary values from column k of the matrix.

(d) [2 pt] Compare the BSP cost of this communication in stage k to that of the LU decomposition. You may give the result as an approximation.