Mastermath midterm examination
Parallel Algorithms.

Teacher: Rob H. Bisseling, Utrecht University

October 24, 2018

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

1. (a) [5 pt] What is an h-relation in the BSP model?

    (b) [5 pt] What is the cost of an h-relation?

2. An *artificial neural network* has several layers, each consisting of artificial neurons and having connections to the previous and next layer. Consider two layers: an input layer represented by a vector $\mathbf{x}$ of length n, where x_i is the activation value at neuron i of this layer, and an output layer represented by a vector $\mathbf{y}$ of length n.

    (a) [5 pt] Assume that the input and output vectors are distributed by the block distribution over p processors with $n \bmod p = 0$. Assume that the output value y_i is computed by

    $$y_i = \max(a_i x_{i-1} + b_i x_i + c_i x_{i+1} - d_i, 0),$$

    where the values a_i, b_i, c_i are weights that have been obtained by previous training of the network, and d_i is a bias value, also obtained by training. The weights and the bias are hence represented by vectors of length n, also in the block distribution. Define dummy values $x_{-1} = x_n = 0$, to define the boundary behaviour.

    Give an efficient BSP algorithm for processor $P(s)$ for this computation in the notation we have learned.

    (b) [5 pt] Analyse the BSP cost.

3. Consider the LU decomposition of an $n \times n$ matrix A without pivoting.

(a) [2 pt] In stage k of the sequential LU decomposition without pivoting, all matrix elements a_{ij} with $k < i, j < n$ are updated by the statement $a_{ij} := a_{ij} - a_{ik}a_{kj}$. Determine the total cost of the matrix updates in this algorithm. You may use the formula $\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

(b) [3 pt] Formulate the computation superstep of stage k of the parallel LU decomposition that corresponds to the matrix update. Assume that we use the square block distribution, with $p = M^2$ processors and $n \bmod M = 0$. Use the notation we have learned to express algorithms.

(c) [3 pt] Analyse the computation cost of the matrix updates in the parallel LU decomposition algorithm for the square block distribution.

(d) [2 pt] What is the maximum speedup on p processors that the parallel LU decomposition algorithm can obtain, compared to the sequential algorithm?

4. *Rook pivoting* is a form of pivoting that is more stable than partial pivoting; it is almost as good as complete pivoting, but has a lower cost in finding the pivot. Rook pivoting works as follows: in stage k of the LU decomposition, we search for an element a_{rc} that has maximum absolute value in both its row and its column, among all elements a_{ij} with $i, j \geq k$. This element becomes the pivot element. Then rows r and k are swapped, and columns c and k. The final result is an LU decomposition $PAQ = LU$, where P and Q are permutation matrices.

Now assume that we are at the start of the LU decomposition (i.e., $k = 0$) and that all matrix elements are different, so we do not have to break ties. Assume we have $p = M^2$ processors and $n \bmod M = 0$. Assume the matrix is distributed by the square cyclic distribution.

(a) [5 pt] Design a BSP algorithm that quickly finds a rook pivot. Where possible and useful, alternate between rows and columns to give both equal treament. It is sufficient to express the algorithm in words (so no detailed program text is needed here).

(b) [5 pt] Analyse its BSP cost.