PARALLEL MERGE SORT USING COMPARISON MATRICES. II

Ya. E. Romm

In this paper, merge techniques using comparison matrices (CMs) [1] are modified to handle several ordered one-dimensional arrays. The modifications are similar to m-way merge [2, 3], and differ only by construction and parallelism features that are specific to the use of CMs. The proposed modifications improve the sorting time bounds derived in [1]. The sorting time $T(R)$ (where $R$ is the number of processor elements) is measured by the number of sequentially executed binary comparisons. In particular we have obtained the bounds $T(N^2) = O(\log_m N)$ and $T(N^2) = O(1)$ for sorting a one-dimensional $N$-element array. Modified merging by comparison matrices is used for parallel sorting of an unbounded one-dimensional array given the main memory size. As an application, we consider merge sorting by CMs (M-sort or M-sorting) to generate the characteristic features of an image represented by a coordinate array.

Multiway merging is called m-merging in the particular case when it is applied to $m$ one dimensional arrays ordered by the relation $\leq$,

$$\{a_{i}^{(j)}\}_{i=1}^{n}, \quad a_{i}^{(j)} \leq a_{i+1}^{(j)} \leq a_{n}^{(j)}, \quad i = 1, 2, \ldots, n-1, \quad j = 1, 2, \ldots, m,$$

transforming them into a single ordered array of $nm$ elements. Initially $a_{i}^{(j)}$ in (1) are assumed nonnegative integers, and the order relation is arithmetic inequality. For any $j = \text{const}$ the array $\{a_{i}^{(j)}\}_{i=1}^{n}$ in (1) is called a segment of length $n$. We assume that m-merging always preserves the initial relative order of equal elements, also when these elements originate from different segments in (1). The order of the segments is always specified. For an arbitrary pair of segments in (1) we form the CM

$$
\begin{array}{cccccc}
& a_{0}^{(j)} & a_{1}^{(j)} & \cdots & a_{n}^{(j)} & a_{n+1}^{(j)} \\
\alpha_{0}^{(j)} & a_{0}^{(j)} & a_{1}^{(j)} & \cdots & a_{n}^{(j)} & a_{n+1}^{(j)} \\
\alpha_{1}^{(j)} & a_{1}^{(j)} & a_{1}^{(j)} & \cdots & a_{1}^{(j)} & a_{1}^{(j)} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\alpha_{n}^{(j)} & a_{n}^{(j)} & a_{n}^{(j)} & \cdots & a_{n}^{(j)} & a_{n}^{(j)} \\
\alpha_{n+1}^{(j)} & a_{n+1}^{(j)} & a_{n+1}^{(j)} & \cdots & a_{n+1}^{(j)} & a_{n+1}^{(j)} \\
\end{array}
$$

where $\{a_{i}^{(j)}\}_{i=1}^{n}, \{a_{i}^{(j)}\}_{i=1}^{n}$ are two segments in (1) augmented with the delimiters $\alpha_{0}^{(j)} = -\infty = a_{0}^{(j)}, \alpha_{n+1}^{(j)} = \infty = a_{n+1}^{(j)}$.

$$
\alpha_{k}^{(j)} = \text{sign} (a_{k}^{(j)} - a_{l}^{(j)}) =
\begin{cases}
-1, & a_{k}^{(j)} < a_{l}^{(j)} \\
0, & a_{k}^{(j)} = a_{l}^{(j)}, \quad n+1 \geq l > 0 \leq k \leq n+1, \\
1, & a_{l}^{(j)} < a_{k}^{(j)}.
\end{cases}
$$

The elements \((3)\) that by definition correspond to 

\[
\alpha_{00}^{(l)} = 0 = \alpha_{(n+1)(n+1)}^{(l)}, \quad \alpha_{lk}^{(l)} = 1 = \alpha_{(n+1)k}^{(l)} \quad (k \neq 0, \ l \neq n + 1),
\]

\[
\alpha_{lk}^{(l)} = -1 = \alpha_{(n+1)k}^{(l)} \quad (l \neq 0, \ k \neq n + 1),
\]

for all \(k, l\) except those excluded in \((4)\) form the bordering of the CM \([1]\); the rows and the columns with these elements are called bordering rows and columns. A CM without bordering is denoted \(\text{CM}_0\). Based on \((2)\) we introduce the indexing \(\text{CM}_{(i,j)}^{(l)}\), \(\text{CM}_0^{(l)}\). The values in \((3)\) are called plus, zero, and minus \([1]\), and are respectively denoted by the symbols \(+, 0, -\). The segments \((1)\) participating in the m-merge are called m-merging segments, and the result of the m-merge is called an m-merged segment. Suppose that the segments \((1)\) are ordered left to right by \(j = 1, 2, \ldots, m\), and a CM of the form \((2)\) has been constructed for each segment \(i\) paired with each of the segments \((1)\) in this order:

\[
\text{CM}_{(i1)}, \text{CM}_{(i2)}^{(l)}, \ldots, \text{CM}_{(im)}^{(l)},
\]

The segments \(\{a_{(i,j)}^{(l)}\}_{l=1}^{m}\) that are fixed for \((5)\) in turn order the collection of matrices \((5)\) from top down by \(i = 1, 2, \ldots, m\). The "insertion" of each element from any of the m-merging segments is defined by these collections of matrices uniquely and independently of other insertions. Thus, if the m-merged segment is denoted \(\{c_{(i,j)}^{(l)}\}_{l=1}^{m}\), and the subset of its elements originating from the m-merging segment \(\{a_{(i,j)}^{(l)}\}_{l=1}^{m}\), is denoted \(\{c_{(i,j)}^{(l)}\}_{l=1}^{m}\), where \(c_{(i,j)}^{(l)} = a_{(i,j)}^{(l)}\), when

\[
k_i = l + j_1 + j_2 + \ldots + j_{i-1} + j_{i+1} + j_{i+2} + \ldots + j_n.
\]

Here \(l = \text{const}\) is the row index in each \(\text{CM}_{(i,j)}^{(l)}\) from \((5)\) for \(i = \text{const} \neq j, j\), is the index of the column in \(\text{CM}_{(i,j)}^{(l)}\), \(i \neq r\), in which the \(l\)-th row element forms a sign change (SGNCH) \([1]\) with the element of the \((j, + 1)\)-th column. Initially, SGNCH is defined as a pair of elements in a row of \(\text{CM}_{(i,j)}^{(l)}\) from \((2)\) such that

\[
\alpha_{lk}^{(l)} < 0, \quad \alpha_{lk+1}^{(l)} > 0, \quad \text{if} \quad j > i,
\]

\[
\alpha_{lk}^{(l)} \leq 0, \quad \alpha_{lk+1}^{(l)} > 0, \quad \text{if} \quad j \leq i,
\]

where for \(i \neq j\) the indices \(l, k\) may take any values as long as \(n + 1 \geq l \geq 0 \leq k \leq n + 1, 1 \leq i \leq m \geq j \geq 1\). Definition \((7)\) differs from the definition of SGNCH in \([1]\) for 2-merge because, given relationship \((6)\), we must preserve the relative order of equal elements in the m-merged segment. Relationship \((6)\), which we call insertion in what follows, is a trivial generalization of the insertion relation for 2-merge \([1]\), and it is easily verified directly.