Odd-Even Exchange Sorting

Let A_0 be an array with N elements, indexed $A_0[0], \ldots A_0[N-1]$ where each element is 0 or 1. For simplicity, I'll assume that N is even – this reduces the number of cases in the arguments below. I believe that the proof when N is odd is similar. For $0 < i \leq N$, let

$$\begin{array}{rcll} A_{i+1}[2j] & = & \min(A_i[2j], A_i[2j+1]) & \text{, if } i \text{ is even, and } 0 \leq j < N/2 \\ & = & \max(A_i[2j-1], A_i[2j+1]) & \text{, if } i \text{ is odd, and } 0 < j < N/2 \\ & = & A_i[0] & \text{, if } i \text{ is odd, and } j = 0 \\ \\ A_{i+1}[2j+1] & = & \max(A_i[2j], A_i[2j+1]) & \text{, if } i \text{ is even, and } 0 \leq j < N/2 \\ & = & \min(A_i[2j+1], A_i[2j+2]) & \text{, if } i \text{ is odd, and } 0 \leq j < N/2 - 1 \\ & = & \min(A_i[2j-1], A_i[2j+1]) & \text{, if } i \text{ is odd, and } j = 0 \end{array}$$

Our claim is that A_N is sorted.

First I tried some examples that I thought might be indicative of cases that take the most number of steps to sort. Figure 1 shows these examples. Consider the number of steps it takes to get all of the 0s into the correct positions; if the 0s are in the correct places, the 1s will be as well. Looking at the figures, 0s move along diagonal paths (upward, to the right) to reach their destination.

Let

$$\begin{array}{rcll} z_{\mathit{before}}(A_i, k) & = & |j|(A_i[j] = 0) \wedge j < k| & \text{, number of 0s before } A_i[k] \\ z_{\mathit{after}}(A_i, k) & = & |j|(A_i[j] = 0) \wedge j > k| & \text{, number of 0s after } A_i[k] \end{array}$$

If $A_i[k]$ is 0, then $z(A_i, k)$ gives the location where this 0 should be stored in the sorted array. From the examples, it appears that a 0 in location k of A_i must reach location $z_{\mathit{before}}(A_i, k)$ $z_{\mathit{after}}(A_i, k)$ or $z_{\mathit{after}}(A_i, k) + 1$ cycles before the end. Whether it's $z_{\mathit{after}}(A_i, k)$ cycles before or $z_{\mathit{after}}(A_i, k) + 1$ depends on whether the final location is even or odd as described below.

One more observation: For $0 < i \leq N$, we'll say that a 0 moved into $A_i[k]$ if $(A_{i-1}[k] = 1) \wedge (A_i[k] = 0)$. For any $0 < i \leq N$ and $0 \leq k < N$, if a 0 moved into $A_i[k]$, then $i + k$ is odd.

I'll now combine these observations into a conjecture. The idea is to figure out a bound on k when $A_i[k] = 0$ based on i and $z_{\mathit{before}}(A_i, k)$ and $z_{\mathit{after}}(A_i, k)$. The *last* 0 must reach its final location by A_N. This means that a 0 in location k of A_i must reach location $z_{\mathit{before}}(A_i, k)$ by time (iteration of the i loop) $N - z_{\mathit{after}}(A_i, k)$.

If location $z_{\mathit{before}}(A_i, k)$ is even, the zero can only move to its final location on an odd step, and if location $z_{\mathit{before}}(A_i, k)$ is odd, the zero can only move to this location on an even step. In other words, if $z_{\mathit{before}}(A_i, k)$ and $N - z_{\mathit{after}}(A_i, k)$ are both even (or both odd), then the deadline for the 0 arriving at its final location is one iteration of the i loop earlier. If $z_{\mathit{before}}(A_i, k)$ and $N - z_{\mathit{after}}(A_i, k)$ are both even (or both odd) then $N - (z_{\mathit{before}}(A_i, k) + z_{\mathit{before}}(A_i, k)$ is even. By the assumption that N is even, this means that $z_{\mathit{before}}(A_i, k) + z_{\mathit{before}}(A_i, k$ is even. Let N_0 be the number of 0s in the array, and note that if $A_i[k] = 0$, then $z_{\mathit{before}}(A_i, k) + z_{\mathit{after}}(A_i, k) = N_0 - 1$. Thus, $z_{\mathit{before}}(A_i, k) + z_{\mathit{before}}(A_i, k)$ is even iff N_0 is odd. This means that a 0 in location k of A_i must reach location $z_{\mathit{before}}(A_i, k)$ by time $N - z_{\mathit{after}}(A_i, k)$ if N_0 is even, and by time $N - z_{\mathit{after}}(A_i, k) - 1$ if N_0 is odd. This gives us a simple formula for when a 0 in position k of A_i must reach position $z_{\mathit{before}}(A_i, k)$:

A 0 in location k of A_i must reach location $z_{\mathit{before}}(A_i, k)$ by time

$$N - z_{\mathit{after}}(A_i, k) - (N_0 \bmod 2)$$

Finally, I'll propose that if a 0 is to make it to its final location by step s, it can be at most $s - i$ locations away on step i (for $i < s$). This condition is clearly necessary. Now, I'll try to prove by induction that it is also sufficient.

Lemma: For all $0 \leq i \leq N$ and all $0 \leq k < N$ if $A_i[k] = 0$, then

$$\begin{array}{lrcl} & k - z_{\mathit{before}}(A_i, k) + i & \leq & N - z_{\mathit{after}}(A_i, k) - (N_0 \bmod 2) \\ \text{or} & k & = & z_{\mathit{before}}(A_i, k) \end{array}$$

```
  i 0 1 2 3 4 5 6 7 8 9 10
k + -----------------------
0 | 1 1 1 1 1 0 0 0 0 0 0
1 | 1 1 1 1 0 1 0 0 0 0 0
2 | 1 1 1 0 1 0 1 0 0 0 0
3 | 1 1 0 1 0 1 0 1 0 0 0
4 | 1 0 1 0 1 0 1 0 1 0 0
5 | 0 1 0 1 0 1 0 1 0 1 1
6 | 0 0 1 0 1 0 1 0 1 1 1
7 | 0 0 0 1 0 1 0 1 1 1 1
8 | 0 0 0 0 1 0 1 1 1 1 1
9 | 0 0 0 0 0 1 1 1 1 1 1
```

```
  i 0 1 2 3 4 5 6 7 8
k +------------------
0 | 1 1 1 1 1 0 0 0 0
1 | 1 1 1 1 0 1 0 0 0
2 | 1 1 1 0 1 0 1 0 0
3 | 1 1 0 1 0 1 0 1 0
4 | 0 0 1 0 1 0 1 0 1
5 | 0 0 0 1 0 1 0 1 1
6 | 0 0 0 0 1 0 1 1 1
7 | 0 0 0 0 0 1 1 1 1
```

```
  i 0 1 2 3 4 5 6 7 8 9 10
k + -----------------------
0 | 1 1 1 1 1 1 1 1 1 0 0
1 | 1 1 1 1 1 1 1 1 0 1 0
2 | 1 1 1 1 1 1 1 0 1 0 1
3 | 1 1 1 1 1 1 0 1 0 1 1
4 | 1 1 1 1 1 0 1 0 1 1 1
5 | 1 1 1 1 0 1 0 1 1 1 1
6 | 1 1 1 0 1 0 1 1 1 1 1
7 | 1 1 0 1 0 1 1 1 1 1 1
8 | 0 0 1 0 1 1 1 1 1 1 1
9 | 0 0 0 1 1 1 1 1 1 1 1
```

```
In these diagrams, column i is A_i.
Row k of column k is A_i[k].
Note how 0s move upard along diagonals.
The 0 that settles in row k is there by
   i=N-z_after(A_i, k).
If a 0 settles in row k with k even,
it must arrive on an odd-numbered cycle.
If k is odd, then the 0 must arrive on
an even numbered cycle.
```

Figure 1: Some examples of odd-even exchange sort

Note: The left side of the first inequality is the earliest step that the 0 at location k in A_i can reach location $z_{before}(A_i, k)$. The right side of the inequality is the time by which it "needs" to be there based on observations from the examples in Figure **??**. Thus, the first condition says that the 0 can make it to its destination on time if it moves one position in each time step. The second condition says that the 0 has already made it to its final location.
Proof: by induction on i.

Base case, $i = 0$:
Choose k such that $A_0[k] = 0$.

1. $k < N - z_{after}(A_0, k)$, there are at least $z_{after}(A_0, k)$ positions in A_i after k
2. $z_{before}(A_i, k) \geq 0$, definition of z_{before}
3. $N_0 \bmod 2 \geq 0$, it's either 0 or 1
4. $k - z_{before}(A_0, k) + 0 \leq N - z_{after}(A_0, k) - (N_0 \bmod 2)$, 1, 2, and 3

Induction step, $0 < i \leq N$. Choose k such that $A_i[k] = 0$. I'll consider four cases based on the even or oddness of i and k.

i is even and k is even: If $k = 0$ then $z_{before}(A_i, k) = 0$ and the 0 is in its final location. Otherwise, A_i was obtained from A_{i-1} by performing a compare and swap between locations $2j + 1$ and $2j + 2$ for $0 \leq j < N/2 - 1$. Because k is even, there is some j such that $k = 2j + 2$. This means that

$$A_{i-1}[k-1] = A_{i-1}[k] = A_i[i-1] = A_i[k] = 0 z_{before}(A_{i-1}, k-1) = z_{before}(A_i, k-1) = z_{before}(A_i, k) - 1 = z_{before}(A_{i-1}$$

From the induction hypothesis for $i - 1$ and $k - 1$:

$$\begin{aligned}
(k-1) - z_{before}(A_{i-1}, k-1) + (i-1) &\leq N - z_{after}(A_{i-1}, k-1) - (N_0 \bmod 2) \\
k - z_{before}(A_i, k-1) + i - 2 &\leq N - z_{after}(A_i, k-1) - (N_0 \bmod 2) \\
k - (z_{before}(A_i, k) - 1) + i - 2 &\leq N - (z_{after}(A_i, k) + 1) - (N_0 \bmod 2) \\
k - z_{before}(A_i, k) + i &\leq N - z_{after}(A_i, k) - (N_0 \bmod 2)
\end{aligned}$$

which shows that the induction hypothesis continues to hold.

i is even and k is odd: If $A_{i-1}[k] = 1$, then $A_{i-1}[k+1] = 0$; $z_{before}(A_i, k) = z_{before}(A_{i-1}, k+1)$; and $z_{after}(A_i, k) = z_{after}(A_{i-1}, k+1)$. From the induction hypothesis for $i - 1$ and $k + 1$:

$$\begin{aligned}
(k+1) - z_{before}(A_{i-1}, k+1) + (i-1) &\leq N - z_{after}(A_{i-1}, k+1) - (N_0 \bmod 2) \\
k - z_{before}(A_i, k) + i &\leq N - z_{after}(A_i, k) - (N_0 \bmod 2)
\end{aligned}$$

which shows that the induction hypothesis continues to hold in the case that $A_{i-1}[k] = 1$.
If $A_{i-1}[k] = 0$, then the compare-and-swap for $A_{i-1}[k]$ and $A_{i-1}[k = 1]$ did not swap its inputs. Thus, $z_{before}(A_i, k) = z_{before}(A_{i-1}, k)$; and $z_{after}(A_i, k) = z_{after}(A_i, k+1)$. We need to show that we have "one to spare" in the inequality for the induction hypothesis at $i - 1$ and k. Here we exploit the details about oddness and evenness. From the induction hypothesis, we get:

$$\begin{aligned}
& k - z_{before}(A_{i-1}, k) + (i-1) &\leq\; & N - z_{after}(A_{i-1}, k) - (N_0 \bmod 2) \\
\equiv\; & k + i - 1 &\leq\; & N + z_{before}(A_{i-1}, k) - z_{after}(A_{i-1}, k) - (N_0 \bmod 2)
\end{aligned}$$

Because i is even and k is odd, $k + i - 1$ is even. Because $z_{before}(A_{i-1}, k) + z_{after}(A_{i-1}, k) = N_0 - 1$ and N is even, $N + z_{before}(A_{i-1}, k) - z_{after}(A_{i-1}, k) - (N_0 \bmod 2)$ is odd. Therefore,

$$k + i - 1 \quad \neq \quad N + z_{before}(A_{i-1}, k) - z_{after}(A_{i-1}, k) - (N_0 \bmod 2)$$

and we conclude

$$\begin{aligned}
& k + i - 1 &<\; & N + z_{before}(A_{i-1}, k) - z_{after}(A_{i-1}, k) - (N_0 \bmod 2) \\
\equiv\; & k - z_{before}(A_{i-1}, k) + (i-1) &<\; & N - z_{after}(A_{i-1}, k) - (N_0 \bmod 2) \\
\equiv\; & k - z_{before}(A_i, k) + i &\leq\; & N - z_{after}(A_i, k) - (N_0 \bmod 2)
\end{aligned}$$

which shows that the induction hypothesis continues to hold in the case that $A_{i-1}[k] = 0$.

Combining these two cases shows that the induction hypothesis continues to hold in the case that i is even and k is odd.

i is odd and k is even: This is very similar to the case where i is even and k is odd. You can work out the details if you want some practice with proofs of this kind.

i is odd and k is odd: This is very similar to the case where i is even and k is even, but we don't have to consider a special case for $k = 0$. Again, feel free to work out the details if you're interested.

□

From the lemma, it immediately follows that for all $0 \leq k < N$, if $A_N[k] = 0$ then $k = z_{before}(A_N, k)$. In other words, if $A_N[k]$ is zero, then for all $0 \leq j < k$, $A[j] = 0$. This shows that A_N is sorted as claimed.

Exercises

1. Show that $N - 1$ steps are not sufficient to sort all inputs consisting of 0s and 1s.
2. I assumed that N is even. What changes are needed to the proof if N is odd?
3. There must be a simpler proof. I'll check Knuth Vol. 3. Let me know if you find a better/simpler argument.