SHELLSORT

Shell-sort is also called <u>sorting by diminishing increment</u>. This method is motivated by the following observation.

> For an array of size n, elements need to travel, on an average, a distance of $n/3$ positions in order to reach their final destination in the final sorted order.

The observation above suggests that progress towards the final sorted order will be quicker if elements are compared and moved initially over longer rather than shorter distances. This would hopefully "place" each element closer to its final destination much faster.

Example: suppose we have to sort the following 8-element array into an ascending sequence.

80	70	60	50	40	30	20	10

Shell-sort with the diminishing increments 4, 2, 1 proceeds in the following way.

80	70	60	50	40	30	20	10

Iteration 1
compare elements which are apart by a distance of 4

Sort 4 chains of length 2, namely, (A[0], A[4]), (A[1], A[5]), (A[2], A[6]), (A[3], A[7])

40	30	20	10	80	70	60	50

contains 4 sorted chains of 2 elements each

Iteration 2
compare elements which are apart by a distance of 2

Sort 2 chains of length 4, namely (A[0], A[2], A[4], A[6]) and (A[1], A[3], A[5], A[7]) (using an appropriate sorting method)

20	10	40	30	60	50	80	70

contains 2 sorted chains of 4 elements each, namely 20, 40, 60, 80; 10, 30, 50, 70

Iteration 3
compare elements which are apart by a distance of 1

Sort 1 chain of 8 elements A[0], ..., A[7] using an appropriate sorting method

10	20	30	40	50	60	70	80

sorted sequence

- The chains that result in successive rounds are sorted and hence the overall sequence gets increasingly partially sorted as we complete the rounds. Insertion sort and bubble sort are preferred as the sorting methods in the preliminary and final rounds. Insertion sort is the method of choice since it does not rely heavily on exchanges.

- The last round has to necessarily have the distance of comparison = 1.

- With the distance sequence

$$\frac{n}{2}, \frac{n}{4}, \ldots, 4, 2, 1$$

the worst case computational complexity is shown to be $O(n^2)$.

- With the distance sequence

$$2^m - 1,\ 2^{m-1} - 1,\ \ldots,\ 15,\ 7,\ 3,\ 1,$$

suggested by Hibbard, the worst case complexity is shown to be $O(n\sqrt{n})$ since it minimizes the number of times the same two elements are compared in the rounds.