Algorithms and Complexity 2003
Analysis of InsertionSort and BubbleSort

K. Scheffler

August 8, 2003

1 InsertionSort

1.1 Best case analysis

The best case is when the list is already sorted. In this case, there is only one comparison per iteration through the outer loop, giving a total of $N-1$ comparisons.

Thus $B(N) = O(N)$.

1.2 Worst case analysis

The worst case is when there are the maximum number of comparisons for each of the $N-1$ iterations through the outer loop. This is i comparisons for the ith iteration:

$$W(N) = \sum_{i=1}^{N-1} i = (N-1)N/2 \tag{1}$$

Thus $W(N) = O(N^2)$.

1.3 Average case analysis

1.3.1 Inserting the ith element

For the average case analysis, we first find the average number of comparisons required to insert the ith element into the list (remember that before the ith element is inserted there are already i elements in the list; after the insertion there are $i+1$ elements). There are $i+1$ possible positions where the ith element can be inserted: we make the assumption that all of these are equally likely. (As usual this assumption is only an approximation.)

Consider the number of comparisons that are done if the element is inserted in position $i+1, i, i-1, \ldots, 2$: for these locations there are $1, 2, 3, \ldots, i$ comparisons respectively. Further, if the element is inserted in position 1 (all the way through to the top of the list), there are also i comparisons (because we run out of elements to compare it with). So when we average over the $i+1$ possibilities we get:

$$
\begin{aligned}
A(i) &= \frac{1}{i+1}\left[\left(\sum_{p=1}^{i} p\right) + i\right] && (2)\\
&= \frac{1}{i+1}\left[\frac{i(i+1)}{2} + i\right] && (3)\\
&= i/2 + i/(i+1) && (4)\\
&= i/2 + 1 - 1/(i+1) && (5)
\end{aligned}
$$

1.3.2 Inserting all elements

Equation 5 gives the average number of comparisons for inserting the ith element. Adding this up over all elements gives the average number of comparisons required to sort the entire list:

$$
\begin{aligned}
A(N) &= \sum_{i=1}^{N-1} A_i && (6)\\
&= \sum_{i=1}^{N-1} (i/2 + 1 - 1/(i+1)) && (7)\\
&= \sum_{i=1}^{N-1} i/2 + \sum_{i=1}^{N-1} 1 + \sum_{i=1}^{N-1} 1/(i+1) && (8)
\end{aligned}
$$

Note that $\sum_{i=1}^{N-1} 1/(i+1) = \sum_{i=2}^{N} 1/i = \sum_{i=1}^{N} (1/i) - 1 \approx \ln N - 1$. Thus:

$$
\begin{aligned}
A(N) &\approx \frac{1}{2}(N-1)N/2 + (N-1) - (\ln N - 1) && (9)\\
&= (N^2 + 3N - 4)/4 - \ln N - 1 && (10)\\
&\approx N^2/4 && (11)
\end{aligned}
$$

Thus $A(N) = O(N^2)$.

2 BubbleSort

2.1 Best case analysis

The best case is when the list is already sorted. In this case, there is only one iteration through the outer loop, with $N-1$ comparisons.

Thus $B(N) = O(N)$.

2.2 Worst case analysis

The worst case is when we have the maximum number $(N-1)$ iterations through the outer loop. The number of comparisons per iteration starts at $N-1$ for the first iteration, with each subsequent iteration having one fewer comparison:

$$\begin{aligned} W(N) &= (N-1) + (N-2) + (N-3) + \ldots + 1 && (12) \\ &= \sum_{i=1}^{N-1} i && (13) \\ &= (N-1)N/2. && (14) \end{aligned}$$

Thus $W(N) = O(N^2)$.

2.3 Average case analysis

For the average case analysis, we need to consider the $N-1$ different cases where there are i iterations through the outer loop, with i ranging from 1 to $N-1$. We assume that all of these cases are equally likely. (Note that this assumption is fairly crude, but it gives an approximation which is "good enough".)

Averaging over the $N-1$ cases, we get the average case complexity:

$$A(N) = \frac{1}{N-1} \sum_{i=1}^{N-1} C(i) \qquad (15)$$

where $C(i)$ is the number of comparisons done in the first i iterations through the outer loop. (Note an error in the text book (p.66): they state this definition for $C(i)$, but then give the equations as if the definition was that $C(i)$ is the number of comparisons done in the first $N-i$ iterations. In these notes we stick with the original definition and obtain equations that are different from those in the book. The final answer is the same for both derivations, which it has to be since both derivations are correct.)

Now find an expression for $C(i)$:

$$\begin{aligned} C(i) &= \sum_{j=N-1}^{N-i} j && (16) \\ &= \sum_{j=1}^{N-1} j - \sum_{j=1}^{N-i} j && (17) \\ &= (N-1)N/2 - (N-i-1)(N-i)/2 && (18) \\ &= iN - i/2 - i^2/2 && (19) \end{aligned}$$

Substituting this into equation 15 gives:

$$
\begin{aligned}
A(N) &= \frac{1}{N-1}\sum_{i=1}^{N-1}(iN - i/2 - i^2/2) && (20)\\
&= \frac{N(N-1)N}{(N-1)2} - \frac{(N-1)N}{2(N-1)2} - \frac{(N-1)N(2N-1)}{2(N-1)6} && (21)\\
&= N^2/2 - N/4 - (2N^2 - N)/12 && (22)\\
&= N^2/3 - N/6 && (23)
\end{aligned}
$$

Thus $A(N) = O(N^2)$.