

Average Case Analysis

	InpType1			InpType2			InpType3			InpType4		
	<i>n</i> =100	<i>n</i> =1000	<i>n</i> =10000	<i>n</i> =100	<i>n</i> =1000	<i>n</i> =10000	<i>n</i> =100	<i>n</i> =1000	<i>n</i> =10000	<i>n</i> =100	<i>n</i> =1000	<i>n</i> =10000
Version 1	0.000532 2456359 863281	0.00561 9907379 15039	0.07174 2010116 57715	0.000598 52600097 65625	0.00554 2564392 089844	0.06271 1238861 08398	0.000344 61021423 339846	0.00578 6037445 06836	0.06531 8107604 98047	0.000423 28834533 69141	0.006262 3500823 97461	0.06993 8421249 38965
Version 2	0.000642 7288055 419922	0.00882 2679519 65332	0.10002 3221969 60449	0.000670 48072814 9414	0.01000 9765625	0.09353 2466888 42773	0.000807 04689025 87891	0.00822 9494094 848633	0.09975 4476547 2412	0.000998 83079528 80859	0.008401 6799926 7578	0.09537 6443862 91504
Version 3	0.000358 9153289 794922	0.00604 0620803 833008	0.07883 7776184 08203	0.000361 82403564 453123	0.00697 8750228 881836	0.06586 7853164 67285	0.000463 96255493 16406	0.00564 0888214 111328	0.07219 4766998 29101	0.000373 26812744 140623	0.005777 6927947 99805	0.07312 4027252 19727
Version 4	0.000523 5671997 070312	0.01066 5607452 392579	0.07626 5621185 30274	0.000449 84817504 88281	0.00747 0846176 147461	0.07334 8426818 84765	0.000495 67222595 21484	0.00699 6440887 451172	0.07557 3968887 3291	0.000459 62333679 19922	0.007354 6886444 091795	0.07684 2641830 44434

Comments:

In our program, we have 4 different versions of quicksort algorithm, each using a different method of implementing quick sort. These methods are:

Version1: Traditional quick sort algorithm is used. Pivot is chosen as the first element in the list in each iteration.

Version2: Pivot is chosen as a random element of the list in each iteration. After this, quick sort is implemented as usual.

Version3: Lists that are handed off to be quick sorted are randomly permuted before each iteration. Afterward, the pivot is chosen as the first element in the lists.

Version4: Version1's deterministic algorithm is implemented but the pivot is chosen according to the median of three method. Its merits and advantages will be explained further down in the comments.

After running our code successfully and inspecting and comparing the different execution times for each scenario, we have found the following results:

We have observed that for both average and worst case inputs that when the input size is small(100) rather than large(10000), the execution times between the different versions of quick sort and input types were minuscule. We know that for the traditional version of quick sort(Version1), the worst case complexity is $O(n^2)$ and average case complexity is $O(n * \log n)$. Different methods used in the other versions(Version2, Version3, and Version4) are a means to improve the worst case complexity and approach it to $O(n * \log n)$. Hence, they don't provide much improvement in the average run time of the algorithms. So, the complexities of quick sort version2, version3, and version4 remain as $O(n * \log n)$. This fact can easily be observed in the execution times of different scenarios for the average case. Changing all the parameters of our program except for input size which would naturally impact the execution time(input type and quick sort version), didn't provide much difference in the execution time which was overall quite small.

As the input lists are generated randomly and are sent to the program unsorted(for average case), methods used in version2, version3 and version4 for randomizing input data don't have much effect on the run time of the algorithms, unlike worst case where the input lists are sorted, hence adding randomness has a great effect on the runtime of the algorithm. Although version3 seems to outperform the other versions in most cases, runtimes are too close together to make a distinct assumption which is expected since their complexities are the same for the average case($O(n * \log n)$). Even when the input size is 10000, although naturally higher, we don't see such a big leap in the execution time of the algorithms.

Worst Case Analysis

	InpType1			InpType2			InpType3			InpType4		
	$n=100$	$n=1000$	$n=10000$	$n=100$	$n=1000$	$n=10000$	$n=100$	$n=1000$	$n=10000$	$n=100$	$n=1000$	$n=10000$
V e r 1	0.001077 4135589 59961	0.11056 2324523 92578	10.4811 1295700 0732	0.000895 0233459 472656	0.103296 5183258 0566	9.02038 3119583 13	0.000762 46261596 67969	0.08428 0967712 40234	6.53416 4905548 096	0.000372 40982055 66406	0.007115 1256561 2793	0.06500 7686614 99023
V e r 2	0.000628 4713745 117188	0.00854 1107177 734375	0.09211 8263244 6289	0.000700 7122039 794922	0.008261 9190216 06445	0.09250 9746551 51367	0.000564 57519531 25	0.00858 9744567 871094	0.08280 3726196 28906	0.000667 09518432 61719	0.009233 9515686 03516	0.09041 9769287 10938
V e r 3	0.000500 9174346 923828	0.00573 5635757 446289	0.06955 4567337 03613	0.000378 3702850 341797	0.006820 4402923 583984	0.06762 5999450 6836	0.000381 70814514 160156	0.00561 4519119 262695	0.06774 9977111 8164	0.000711 44104003 90625	0.005428 0757904 052734	0.07796 6451644 89746
V e r 4	0.000322 8187561 035156	0.00524 4493484 49707	0.05369 5201873 7793	0.000400 5432128 90625	0.005234 2414855 95703	0.05366 8737411 49902	0.000363 82675170 89844	0.00503 8022994 995117	0.05227 4465560 913086	0.000449 18060302 734375	0.007045 7458496 09375	0.07697 4868774 41406

Comments:

We know that for the traditional version of quick sort(Version1), the worst case complexity is $O(n^2)$. Different methods used in the other versions(Version2, Version3, and Version4) are a means to improve the worst case complexity and approach it to $O(n * \log n)$ which is the best/average case.

Worst case input lists, unlike average case input lists, are already sorted before they are placed in the algorithm to be quick sorted. This makes the randomization used in version2, version3, and version4 of quicksort have a massive effect on the runtime as we know that already sorted lists are the worst type of input for quick sort algorithms.

By observing the execution times of different scenarios, unlike the average case, we could see big changes between the different scenarios. Similar to the average case, when the input size was small(100) rather than big(10000), the execution times were smaller and didn't vary as much. However, in all scenarios, the traditional version of quick sort(version1) performed the worst and had the highest execution time. This is because, in the traditional version of quick sort(version1), the first element in the list is always chosen to be the pivot, hence in each iteration, the next list to be quick sorted always has a maximum length.

In version 2, by choosing a random pivot, this problem is eliminated by placing the pivot in its proper position in the list hence in the next iterations, the list will likely not be as big.

In version3, although the pivot is chosen to be the first element in each iteration like the traditional quick sort algorithm(version1), as the list is permuted prior to being sorted, it is no longer the worst case type of input and performs like the average case.

In version 4, the median of three method is used. In this method, median of the first, middle and last element of the list is chosen as the pivot. These elements are sorted before being placed back in the list. This strategy also eliminates the problem of maximum list size in the next calls of quick sort method and sorting of these three elements provides better picks of pivot in the next runs of the algorithm while keeping the list relatively sorted.

The comparatively poor execution of version1 can be seen by observing the runtimes. For size 100, this difference was not as big but still apparent. As the input sizes got larger, the execution time difference between version1 and the others got bigger and for input size 10000, the difference was massive.

Between version2, version3, and version4, although there isn't much difference since they all approach complexity $O(n * \log n)$, version4(median of three method) mostly performed the best for the reasons presented above. After version4, version3 performed mostly better than version2.

Another difference we can observe between execution times comes from the change in input types. As we go from input type 1 to 4, the range of values that the list elements can take decreases. In fact for input type 4, all of the list elements are 1 independent of size. Hence, as we go from input type 1 to 4, the odds of having more than 1 of the same element in a list increases. When two elements are compared, the possibility of them being equal also increases. Hence, the number of elements to change the positions also decrease. This can result in a quicker run of the algorithm. This notion can easily be observed in the execution times of input 4. Even for quicksort version1 and size 10000, the list is sorted very fast.

