

Counting Sort

Counting sort is a sorting technique based on keys between a specific range. It works by counting the number of objects having distinct key values (kind of hashing). Then doing some arithmetic to calculate the position of each object in the output sequence.

Let us understand it with the help of an example.

For simplicity, consider the data in the range 0 to 9.

Input data: 1, 4, 1, 2, 7, 5, 2

1) Take a count array to store the count of each unique object.

Index:	0	1	2	3	4	5	6	7	8	9
Count:	0	2	2	0	1	1	0	1	0	0

2) Modify the count array such that each element at each index stores the sum of previous counts.

Index:	0	1	2	3	4	5	6	7	8	9
Count:	0	2	4	4	5	6	6	7	7	7

The modified count array indicates the position of each object in the output sequence.

3) Output each object from the input sequence followed by decreasing its count by 1.

Process the input data: 1, 4, 1, 2, 7, 5, 2. Position of 1 is 2. Put data 1 at index 2 in output. Decrease count by 1 to place next data 1 at an index 1 smaller than this index.

Implementation:

```
1 // The main function that sort the given string arr[] in alphabetical order
2 void countSort(char arr[])
3 {
4     // The output character array that will have sorted arr
5     char output[strlen(arr)];
6     // Create a count array to store count of individual characters and initialize count array as 0
7     int count[RANGE + 1], i;
8     memset(count, 0, sizeof(count));
9     // Store count of each character
10    for(i = 0; arr[i]; ++i)
11        ++count[arr[i]];
12    // Change count[i] so that count[i] now contains actual position of this character in output array
13    for (i = 1; i <= RANGE; ++i)
14        count[i] += count[i-1];
15    // Build the output character array
16    for (i = 0; arr[i]; ++i)
17    {
18        output[count[arr[i]]-1] = arr[i];
19        --count[arr[i]];
20    }
21    // Copy the output array to arr, so that arr now contains sorted characters
22    for (i = 0; arr[i]; ++i)
23        arr[i] = output[i];
24 }
25
```

Time Complexity: $O(N + K)$ where N is the number of elements in input array and K is the range of input.

Auxiliary Space: $O(N + K)$

The problem with the previous counting sort was that it could not sort the elements if we have negative numbers in the array because there are no negative array indices. So what we can do is, we can find the minimum element and store count of that minimum element at zero index.

Implementation:

```
1 |
2 | void countSort(vector<int>& arr)
3 | {
4 |     int max = *max_element(arr.begin(), arr.end());
5 |     int min = *min_element(arr.begin(), arr.end());
6 |     int range = max - min + 1;
7 |
8 |     vector<int> count(range), output(arr.size());
9 |     for(int i = 0; i < arr.size(); i++)
10 |         count[arr[i]-min]++;
11 |
12 |     for(int i = 1; i < count.size(); i++)
13 |         count[i] += count[i-1];
14 |
15 |     for(int i = arr.size()-1; i >= 0; i--)
16 |     {
17 |         output[ count[arr[i]-min] - 1 ] = arr[i];
18 |         count[arr[i]-min]--;
19 |     }
20 |
21 |     for(int i=0; i < arr.size(); i++)
22 |         arr[i] = output[i];
23 | }
24 |
```

Important Points:

1. Counting sort is efficient if the range of input data is not significantly greater than the number of objects to be sorted. Consider the situation where the input sequence is between range 1 to 10K and the data is 10, 5, 10K, 5K.
2. It is not a comparison based sorting. It's running time complexity is $O(n)$ with space proportional to the range of data.
3. It is often used as a sub-routine to another sorting algorithm like radix sort.
4. Counting sort uses a partial hashing to count the occurrence of the data object in $O(1)$.
5. Counting sort can be extended to work for negative inputs also.