



**INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH
TECHNOLOGY
SORTIE SORTING**

Tannishk Sharma

* Department of information technology BVCOE, DELHI INDIA

ABSTRACT

Almost all computers regularly sort data. Many different sorting algorithms have been proposed. It is known that no sort algorithm based on key comparisons can sort N keys in less than $O(N \log N)$ operations and that many perform $O(N^2)$ operations in the worst case .

This paper is aimed at proposing a new sortie tree data structure, which can be used for the sorting of data. This algorithm that implements a sortie tree data structure is a non-comparative sorting algorithm.

KEYWORDS: Sorting, Non Comparative Sort, Sortie data structure, sortie sort, msd radix sort , radix sort

INTRODUCTION

Sorting

Let A be a list of n elements A_1, A_2, \dots, A_n in memory. Sorting of A refers to the operation of rearranging the contents of A so that they are either in increasing or in decreasing order (numerically or lexicographically) that is:-

$A_1 > A_2 > A_3 > A_4 > \dots > A_n$

Or

$A_1 < A_2 < A_3 < A_4 < \dots < A_n$

Trees

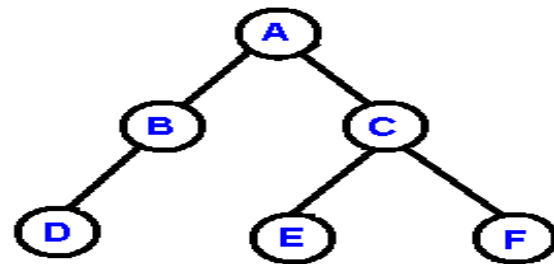
A tree is a connected undirected graph with no simple circuits. A tree cannot contain multiple edges or loops. A tree has a set of vertices and edges.

An undirected graph is a tree if and only if there is a unique simple path between any two of its vertices. In many applications of a tree a particular vertex is designated as the root and we assume direction of each edge is directed away from root.

Some Terminologies used in trees: -

If T is a rooted tree . Let v be a vertex in T other than root, then the parent of v is the unique vertex u such that there is a directed edge from u to v . When u is the parent of v. v is called a child of u . Vertices with same parents are called siblings. A Vertex that has no children is called Leaf node. Vertices that have children are called internal node .

Eg



Here A is the root

B is the parent of D. C is the parent of E and F

E and F are Siblings of each other

A,B,C are internal nodes

D,E,F are leaf nodes

Traversal

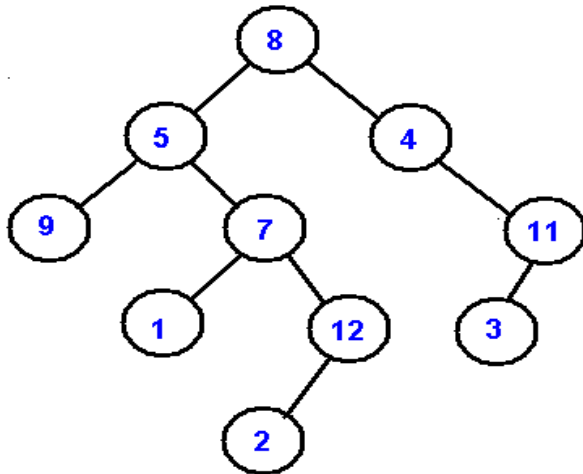
A traversal is the process in which we visit all the nodes of a tree. We divide the traversal algorithms in the following ways :-

- Depth First Traversal
- Breath First Traversal

There are three different types of depth first traversal :-

- Preorder Traversal -- visit the parent first and then left and right children;
- Inorder Traversal – visit the left child , then the parent and the right child
- Postorder Traversal – visit left child, then the right child and then the parent

There is only one kind of breadth-first traversal--the level order traversal. This traversal visits nodes by levels from top to bottom and from left to right.



As an example consider the following tree and its four traversals:

- PreOrder - 8, 5, 9, 7, 1, 12, 2, 4, 11, 3
- InOrder - 9, 5, 1, 7, 2, 12, 8, 4, 3, 11
- PostOrder - 9, 1, 2, 12, 7, 5, 3, 11, 4, 8
- LevelOrder - 8, 5, 4, 9, 7, 11, 1, 12, 3, 2

Radix Sort

It is a fast stable sorting algorithm generally used to sort data on the computer with large memory. It is a non – Comparative Sorting algorithm sorts data with integer keys by grouping keys by the individual digits which share the same significant position and value. Radix Sort is of 2 types MSD (Most significant digit radix sort) this examines the digit from left to right working with most significant digit first. LSD (least significant digit) This examine the digits in the keys in a right-to-left order, working with the least significant digits first. In this algorithm each key is first dropped into first level of bucket corresponding to the value of the rightmost digit. Each bucket preserves the original order of the keys as the keys are dropped into the bucket. There is a one-to-one correspondence between the buckets and the values that can be represented by the rightmost digit. Then the process repeats with the neighboring digits until no more digits are left to process .

SORTIE TREES

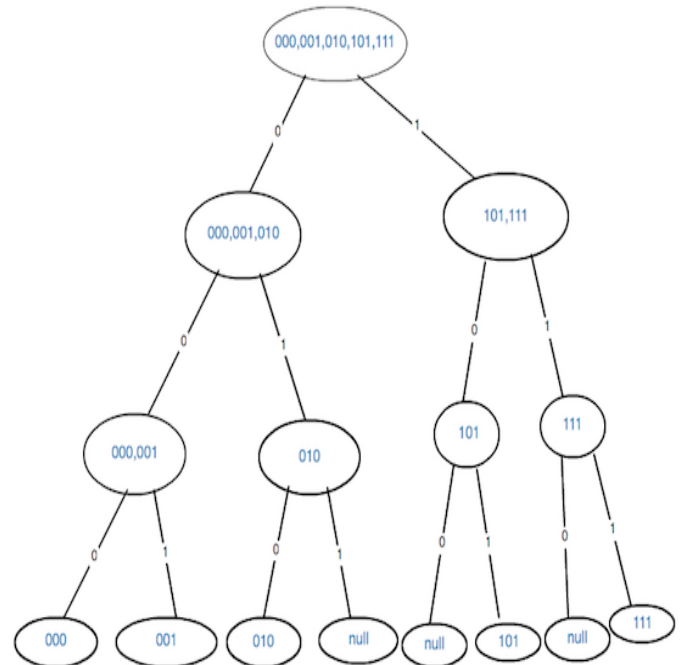
Here I introduce a particular kind of tree called sortie tree . These Types of trees are generally used to sort data with integer keys by grouping keys by the

individual digits which share same significant position and value . This Sorting is not only limited to integers but can also be used to sort Strings and floating point numbers. Sortie trees processes keys from most significant bit then go to the least significant bit.

Properties: -

- Each Node can either has 0 children or number of children is equal to radix (e.g.. 10 in case of decimal system, 2 incase of binary system, 8 in case of octal system)
- Each Internal Node can hold multiple values.
- Each Leaf Node holds only one value
- All the leaf nodes are always at the same level
- Height of the tree is equal to number of maximum digits in the set of keys
- Level order traversal of the tree considering only the leaf node gives the data in sorted order (not considering the null value)

E.g



ALGORITHM FOR INSERTION OF KEYS IN SORTIE TREE

- Check whether the given set of keys have the same base or not. If not then exit
- Check the key with most number of digits. Make the number of digits in other keys equal to the key with most number of digits

- Now the set of key which we need to sort is divided into subsequent set of nodes (one level down in the tree) according to the value of left most digit i.e. the most significant bit
- Numbers starting with 0 goes down by 1st edge to form the 1st node; Starting with 1 goes down by 2nd edge to form the 2nd node and so on. Incase there is no key that has been able to go down the edge then that node down the edge takes a null value
- After most significant bit we move onto a less significant digit and repeat the same process treating each node as a separate tree
- This process continues till there are no more digits left to process.

An Example

Original, unsorted list: -

0,1,10,101,111

As all numbers are of the base 2 hence can be sorted using this algorithm

Key 111 has maximum number of bits. Maximum number of bits is 3.

Now make the number of digits in other keys equal to 3

So

0 – 000

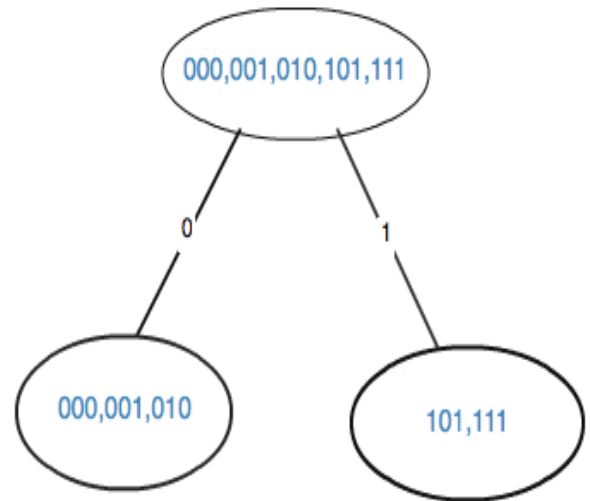
1 – 001

10 – 010

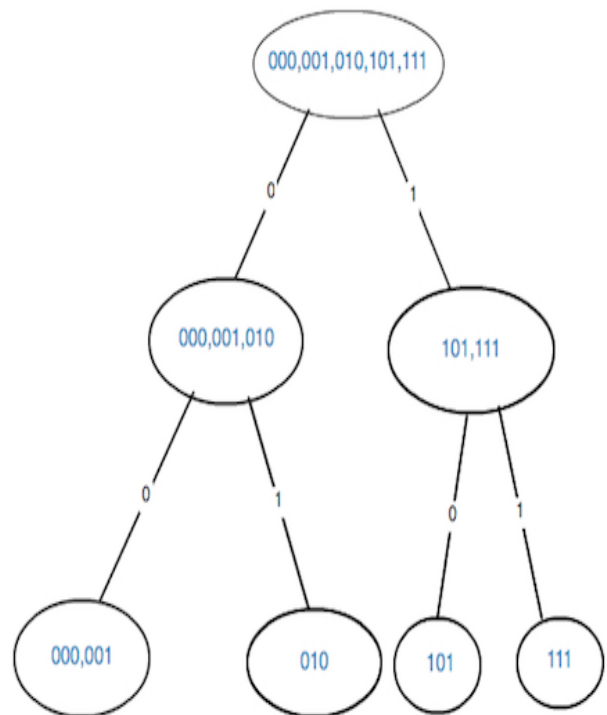
101 – 101

111 – 111

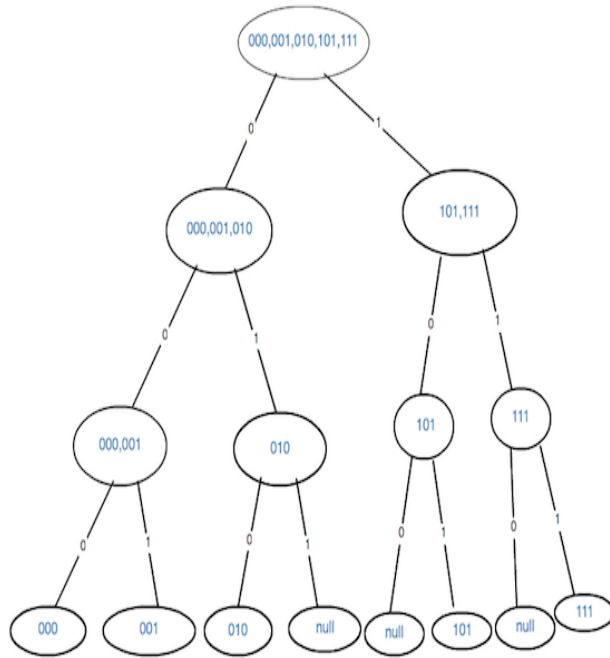
Arranging in Sortie tree according to most significant digit



Arranging a sortie tree according to next digit that is next to most significant bit: -



Arranging a sortie tree according to the next digit:-



As no more digits are left for processing so the above figure shows the final tree. Thereby completing the process of insertion in a sortie tree

ALGORITHM TO FIND DATA IN SORTED ORDER (SORTIE SORT)

- Use the above specified method to create a sortie data structure for the given set of keys
- Do Level order traversal of the tree considering only the leaf node gives the data in sorted order (not considering the null value)

As in the above example: -

Apply Level order traversal of the tree considering only the leaf node gives the data in sorted order (not considering the null value)

We get the data in sorting order: -
 000,001,010,101,111

COMPLEXITY OF SORTIE SORTING ALGORITHM: -

Let a Set of keys consists of n elements A1, A2,... , An is given. Let d donates the radix (e.g. d= 10 for digits, d= 26 for letters, d= 2 for bits) and suppose each item

of Ai consists of s digits. Height of the sortie tree data structure is s. Adding one item to the sortie tree on average takes O(s) time. So adding n items to the tree takes O (n*s) time. To calculate worst case we can suppose s=n. So Worst case is O(n²) . In Best case s=log_d n so best case is O(n log n) .

CONCLUSION

In this paper we have studied and analysed a new sortie data structure. This sortie data structure finds its application in sorting as shown above. The above stated algorithm is a stable non – comparative sorting algorithm. Clearly the given algorithm can do much better job over existing sorting algorithms for a homogenous data set. This algorithm is generally employed on the computers with large memory. The new algorithm proposed can be used by card sorting machines. This algorithm can also be employed in the sorting of words in a dictionary and also to sort records in databases by the primary key value.

REFERENCES

- [1] Kenneth H Rosen , Discrete Mathematics and its applications , 7nd ed. McGraw Hill Education New Delhi
- [2] Seymour Lipchitz , GA Vijayalakshmi Pai, Data Structures , Tata- McGraw Hill Education New Delhi
- [3] PhD. Thesis by Aditya Dev Mishra under the supervision of Dr. Deepak Garg CSED CSE Department Thapar University Patiala-147004 June 2009 on the Topic” Selection of Best Sorting Algorithm for a particular problem.
- [4] <http://www.cs.cmu.edu/~adamchik/15-121/lectures/Trees/trees.html>
- [5] <http://www.csc.villanova.edu/~japaridz/Archive/1300/lect8.1/tsld002.htm>
- [6] http://en.wikipedia.org/wiki/Radix_sort
- [7] Comparison of Sorting Algorithms based on Input Sequences International Journal of Computer Applications (0975 – 8887)Volume 78 – No.14, September 2013
- [8] Formulation and analysis of in-place MSD radix sort algorithms NasirAl-Darwish Journal of Information Science; 31; 467