Homework 3

Name : Vikhyat Dhamija

RUID: 194003013

Net id: vd283

Q1 Develop an implementation of the basic symbol-table API that uses 2-3 trees that are not *necessarily balanced as the underlying data structure*. Allow 3-nodes to lean either way. Hook the new node onto the bottom with a black link when inserting into a 3-node at the bottom.

Solution:

Program name q1.py is a 2-3 Tree implementation (self code–no help from any material except the algorithm learnt in the class)-

As it was not clear what it meant by the unbalanced 2-3 tree, I have implemented the balanced 2-3 tree which was explained by the professor before the Red- Black Tree. (confirmed with Ghadir also)

I have tested it for various inputs. Here is the sample inputs given and the corresponding outputs.:

Note- Input list can be changed for seeing the desired list

#For Random inputs with duplicate values also

#Note that the values are being taken as same as of key(in key -value pair) for simplicity and when the same key is entered twice then the value of the key is updated with the latest value

Input of the Program:

[13,7,13,24,15,4,29,4,20,16,19,1,5,22,5,31,32,2,3,25,27]

Output of the program is as shown:

----------Inorder Traversal----------
The value of the Key : 1
The value of the Value : 1
The value of the Key : 2
The value of the Value : 2
The value of the Key : 3
The value of the Value : 3
The value of the Key : 4
The value of the Value : 4
The value of the Key : 5
The value of the Value : 5
The value of the Key : 7
The value of the Value : 7
The value of the Key : 13
The value of the Value : 13
The value of the Key : 15
The value of the Value : 15
The value of the Key : 16
The value of the Value : 16
The value of the Key : 19
The value of the Value : 19
The value of the Key : 20

```
The value of the Value :  20
The value of the Key :  22
The value of the Value :  22
The value of the Key :  24
The value of the Value :  24
The value of the Key :  25
The value of the Value :  25
The value of the Key :  27
The value of the Value :  27
The value of the Key :  29
The value of the Value :  29
The value of the Key :  31
The value of the Value :  31
The value of the Key :  32
The value of the Value :  32
----------The Level wise Tree----------
The Level  1 :
    [16, None]
The Level  2 :
    [4, None]
    [24, None]
The Level  3 :
    [2, None]
    [13, None]
    [20, None]
    [27, 31]
The Level  4 :
    [1, None]
    [3, None]
    [5, 7]
    [15, None]
    [19, None]
    [22, None]
    [25, None]
    [29, None]
    [32, None]
-------To check the Get value Function----------
Please input the key to get the value5
The value of the Key :  5  is :  5
```

Input of the Program:

[1,2,3,4,5,6,7,8,9,9,9,10,11,12,13,14,15,16,17]

Output of the program is as shown:

```
----------Inorder Traversal----------
The value of the Key :  1
The value of the Value :  1
The value of the Key :  2
The value of the Value :  2
The value of the Key :  3
The value of the Value :  3
The value of the Key :  4
The value of the Value :  4
The value of the Key :  5
```

```
The value of the Value :  5
The value of the Key :  6
The value of the Value :  6
The value of the Key :  7
The value of the Value :  7
The value of the Key :  8
The value of the Value :  8
The value of the Key :  9
The value of the Value :  9
The value of the Key :  10
The value of the Value :  10
The value of the Key :  11
The value of the Value :  11
The value of the Key :  12
The value of the Value :  12
The value of the Key :  13
The value of the Value :  13
The value of the Key :  14
The value of the Value :  14
The value of the Key :  15
The value of the Value :  15
The value of the Key :  16
The value of the Value :  16
The value of the Key :  17
The value of the Value :  17

----------The Level wise Tree----------
The Level  1 :
    [8, None]
The Level  2 :
    [4, None]
    [12, None]
The Level  3 :
    [2, None]
    [6, None]
    [10, None]
    [14, 16]
The Level  4 :
    [1, None]
    [3, None]
    [5, None]
    [7, None]
    [9, None]
    [11, None]
    [13, None]
    [15, None]
    [17, None]
```

#So we can see the balance tree has put the 17 elements above in the 2-3 tree of the sorted array in the 4 Levels where root is considered as level 1 so Log N implementation has been achieved in comparison to BST where it would have been the O(N).

Note 2 lists are there and one can be commented and other can be used to see the above results.

Q2. Run experiments to develop a hypothesis estimating the average path length in a tree built from (i) N-random insertions. (ii) N-sorted insertions?
Solution:

Results from the program execution:

----------For the random elements----------

The tree with size : 2 the average path length is : 0.5
The tree with size : 4 the average path length is : 1.5
The tree with size : 8 the average path length is : 1.75
The tree with size : 16 the average path length is : 3.3125
The tree with size : 32 the average path length is : 3.875
The tree with size : 64 the average path length is : 5.203125
The tree with size : 128 the average path length is : 6.703125
The tree with size : 256 the average path length is : 8.0078125
The tree with size : 512 the average path length is : 9.986328125
The tree with size : 1024 the average path length is : 10.0947265625
The tree with size : 2048 the average path length is : 12.34765625
The tree with size : 4096 the average path length is : 13.510986328125
The tree with size : 8192 the average path length is : 14.7877197265625

----------For the sorted elements----------

The tree with size : 2 the average path length is : 0.5
The tree with size : 4 the average path length is : 1.5
The tree with size : 8 the average path length is : 3.5
The tree with size : 16 the average path length is : 7.5
The tree with size : 32 the average path length is : 15.5
The tree with size : 64 the average path length is : 31.5
The tree with size : 128 the average path length is : 63.5
The tree with size : 256 the average path length is : 127.5
The tree with size : 512 the average path length is : 255.5
The tree with size : 1024 the average path length is : 511.5
The tree with size : 2048 the average path length is : 1023.5
The tree with size : 4096 the average path length is : 2047.5
The tree with size : 8192 the average path length is : 4095.5

From the above , it can be clearly observed that the average internal path length is proportion to Log(N) in case of random insertions and is proportional to N in case of sorted insertions.

Note : - I have implemented the two functions for calculating the internal path length i.e. by using the recursive formula :

A. **Path length(P)=P(Left Subtree)+ P(Right Subtree)+Number of Nodes in a Tree -1**
B. **Using the traversal , we calculated the level of each node starting from 0 of the node and then sum up all the levels of the node to get the path length**

After getting the path length we can get the average path length.

Q3. Write a program that computes the percentage of red nodes in a given red-black tree. Test

program by running at least 100 trials of the experiment of increasing N random keys into an initially empty tree for N=10^4, 10^5 and 10^6 and formulate a hypothesis.

Solution:

For N = 10000 random numbers , 100 trials result:

Trial Number : 0 Red nodes : 2498 Black nodes : 7502 Per of red : 24.98
Trial Number : 1 Red nodes : 2487 Black nodes : 7513 Per of red : 24.87
Trial Number : 2 Red nodes : 2521 Black nodes : 7479 Per of red : 25.21
Trial Number : 3 Red nodes : 2592 Black nodes : 7408 Per of red : 25.919999999999998
Trial Number : 4 Red nodes : 2514 Black nodes : 7486 Per of red : 25.14
Trial Number : 5 Red nodes : 2525 Black nodes : 7475 Per of red : 25.25
Trial Number : 6 Red nodes : 2549 Black nodes : 7451 Per of red : 25.490000000000002
Trial Number : 7 Red nodes : 2500 Black nodes : 7500 Per of red : 25.0
Trial Number : 8 Red nodes : 2549 Black nodes : 7451 Per of red : 25.490000000000002
Trial Number : 9 Red nodes : 2489 Black nodes : 7511 Per of red : 24.89
Trial Number : 10 Red nodes : 2533 Black nodes : 7467 Per of red : 25.330000000000002
Trial Number : 11 Red nodes : 2542 Black nodes : 7458 Per of red : 25.419999999999998
Trial Number : 12 Red nodes : 2538 Black nodes : 7462 Per of red : 25.380000000000003
Trial Number : 13 Red nodes : 2582 Black nodes : 7418 Per of red : 25.82
Trial Number : 14 Red nodes : 2570 Black nodes : 7430 Per of red : 25.7
Trial Number : 15 Red nodes : 2557 Black nodes : 7443 Per of red : 25.569999999999997
Trial Number : 16 Red nodes : 2562 Black nodes : 7438 Per of red : 25.619999999999997
Trial Number : 17 Red nodes : 2515 Black nodes : 7485 Per of red : 25.15
Trial Number : 18 Red nodes : 2539 Black nodes : 7461 Per of red : 25.39
Trial Number : 19 Red nodes : 2614 Black nodes : 7386 Per of red : 26.14
Trial Number : 20 Red nodes : 2571 Black nodes : 7429 Per of red : 25.71
Trial Number : 21 Red nodes : 2538 Black nodes : 7462 Per of red : 25.380000000000003
Trial Number : 22 Red nodes : 2538 Black nodes : 7462 Per of red : 25.380000000000003
Trial Number : 23 Red nodes : 2602 Black nodes : 7398 Per of red : 26.02
Trial Number : 24 Red nodes : 2500 Black nodes : 7500 Per of red : 25.0
Trial Number : 25 Red nodes : 2542 Black nodes : 7458 Per of red : 25.419999999999998
Trial Number : 26 Red nodes : 2475 Black nodes : 7525 Per of red : 24.75
Trial Number : 27 Red nodes : 2580 Black nodes : 7420 Per of red : 25.8
Trial Number : 28 Red nodes : 2535 Black nodes : 7465 Per of red : 25.35
Trial Number : 29 Red nodes : 2577 Black nodes : 7423 Per of red : 25.77
Trial Number : 30 Red nodes : 2537 Black nodes : 7463 Per of red : 25.369999999999997
Trial Number : 31 Red nodes : 2547 Black nodes : 7453 Per of red : 25.47
Trial Number : 32 Red nodes : 2532 Black nodes : 7468 Per of red : 25.319999999999997
Trial Number : 33 Red nodes : 2527 Black nodes : 7473 Per of red : 25.27
Trial Number : 34 Red nodes : 2556 Black nodes : 7444 Per of red : 25.56
Trial Number : 35 Red nodes : 2465 Black nodes : 7535 Per of red : 24.65
Trial Number : 36 Red nodes : 2520 Black nodes : 7480 Per of red : 25.2
Trial Number : 37 Red nodes : 2487 Black nodes : 7513 Per of red : 24.87
Trial Number : 38 Red nodes : 2574 Black nodes : 7426 Per of red : 25.740000000000002
Trial Number : 39 Red nodes : 2519 Black nodes : 7481 Per of red : 25.19
Trial Number : 40 Red nodes : 2521 Black nodes : 7479 Per of red : 25.21
Trial Number : 41 Red nodes : 2594 Black nodes : 7406 Per of red : 25.94
Trial Number : 42 Red nodes : 2551 Black nodes : 7449 Per of red : 25.509999999999998
Trial Number : 43 Red nodes : 2565 Black nodes : 7435 Per of red : 25.650000000000002
Trial Number : 44 Red nodes : 2600 Black nodes : 7400 Per of red : 26.0
Trial Number : 45 Red nodes : 2577 Black nodes : 7423 Per of red : 25.77
Trial Number : 46 Red nodes : 2562 Black nodes : 7438 Per of red : 25.619999999999997
Trial Number : 47 Red nodes : 2509 Black nodes : 7491 Per of red : 25.09
Trial Number : 48 Red nodes : 2559 Black nodes : 7441 Per of red : 25.590000000000003

Trial Number : 49 Red nodes : 2558 Black nodes : 7442 Per of red : 25.580000000000002
Trial Number : 50 Red nodes : 2544 Black nodes : 7456 Per of red : 25.44
Trial Number : 51 Red nodes : 2584 Black nodes : 7416 Per of red : 25.840000000000003
Trial Number : 52 Red nodes : 2513 Black nodes : 7487 Per of red : 25.130000000000003
Trial Number : 53 Red nodes : 2537 Black nodes : 7463 Per of red : 25.369999999999997
Trial Number : 54 Red nodes : 2577 Black nodes : 7423 Per of red : 25.77
Trial Number : 55 Red nodes : 2549 Black nodes : 7451 Per of red : 25.490000000000002
Trial Number : 56 Red nodes : 2538 Black nodes : 7462 Per of red : 25.380000000000003
Trial Number : 57 Red nodes : 2544 Black nodes : 7456 Per of red : 25.44
Trial Number : 58 Red nodes : 2548 Black nodes : 7452 Per of red : 25.480000000000004
Trial Number : 59 Red nodes : 2567 Black nodes : 7433 Per of red : 25.669999999999998
Trial Number : 60 Red nodes : 2568 Black nodes : 7432 Per of red : 25.679999999999996
Trial Number : 61 Red nodes : 2492 Black nodes : 7508 Per of red : 24.92
Trial Number : 62 Red nodes : 2564 Black nodes : 7436 Per of red : 25.64
Trial Number : 63 Red nodes : 2509 Black nodes : 7491 Per of red : 25.09
Trial Number : 64 Red nodes : 2561 Black nodes : 7439 Per of red : 25.61
Trial Number : 65 Red nodes : 2512 Black nodes : 7488 Per of red : 25.119999999999997
Trial Number : 66 Red nodes : 2497 Black nodes : 7503 Per of red : 24.97
Trial Number : 67 Red nodes : 2511 Black nodes : 7489 Per of red : 25.11
Trial Number : 68 Red nodes : 2518 Black nodes : 7482 Per of red : 25.180000000000003
Trial Number : 69 Red nodes : 2544 Black nodes : 7456 Per of red : 25.44
Trial Number : 70 Red nodes : 2495 Black nodes : 7505 Per of red : 24.95
Trial Number : 71 Red nodes : 2531 Black nodes : 7469 Per of red : 25.31
Trial Number : 72 Red nodes : 2498 Black nodes : 7502 Per of red : 24.98
Trial Number : 73 Red nodes : 2546 Black nodes : 7454 Per of red : 25.46
Trial Number : 74 Red nodes : 2556 Black nodes : 7444 Per of red : 25.56
Trial Number : 75 Red nodes : 2566 Black nodes : 7434 Per of red : 25.66
Trial Number : 76 Red nodes : 2533 Black nodes : 7467 Per of red : 25.330000000000002
Trial Number : 77 Red nodes : 2596 Black nodes : 7404 Per of red : 25.96
Trial Number : 78 Red nodes : 2503 Black nodes : 7497 Per of red : 25.03
Trial Number : 79 Red nodes : 2469 Black nodes : 7531 Per of red : 24.69
Trial Number : 80 Red nodes : 2545 Black nodes : 7455 Per of red : 25.45
Trial Number : 81 Red nodes : 2563 Black nodes : 7437 Per of red : 25.629999999999995
Trial Number : 82 Red nodes : 2589 Black nodes : 7411 Per of red : 25.89
Trial Number : 83 Red nodes : 2551 Black nodes : 7449 Per of red : 25.509999999999998
Trial Number : 84 Red nodes : 2503 Black nodes : 7497 Per of red : 25.03
Trial Number : 85 Red nodes : 2550 Black nodes : 7450 Per of red : 25.5
Trial Number : 86 Red nodes : 2580 Black nodes : 7420 Per of red : 25.8
Trial Number : 87 Red nodes : 2595 Black nodes : 7405 Per of red : 25.95
Trial Number : 88 Red nodes : 2540 Black nodes : 7460 Per of red : 25.4
Trial Number : 89 Red nodes : 2541 Black nodes : 7459 Per of red : 25.41
Trial Number : 90 Red nodes : 2488 Black nodes : 7512 Per of red : 24.88
Trial Number : 91 Red nodes : 2533 Black nodes : 7467 Per of red : 25.330000000000002
Trial Number : 92 Red nodes : 2559 Black nodes : 7441 Per of red : 25.590000000000003
Trial Number : 93 Red nodes : 2516 Black nodes : 7484 Per of red : 25.16
Trial Number : 94 Red nodes : 2561 Black nodes : 7439 Per of red : 25.61
Trial Number : 95 Red nodes : 2511 Black nodes : 7489 Per of red : 25.11
Trial Number : 96 Red nodes : 2591 Black nodes : 7409 Per of red : 25.91
Trial Number : 97 Red nodes : 2531 Black nodes : 7469 Per of red : 25.31
Trial Number : 98 Red nodes : 2601 Black nodes : 7399 Per of red : 26.009999999999998
Trial Number : 99 Red nodes : 2519 Black nodes : 7481 Per of red : 25.19

For N = 100000 random numbers , 100 trials result:

Trial Number : 0 Red nodes : 25371 Black nodes : 74629 Per of red : 25.371
Trial Number : 1 Red nodes : 25443 Black nodes : 74557 Per of red : 25.442999999999998
Trial Number : 2 Red nodes : 25419 Black nodes : 74581 Per of red : 25.419000000000004
Trial Number : 3 Red nodes : 25411 Black nodes : 74589 Per of red : 25.411
Trial Number : 4 Red nodes : 25406 Black nodes : 74594 Per of red : 25.406000000000002
Trial Number : 5 Red nodes : 25409 Black nodes : 74591 Per of red : 25.409
Trial Number : 6 Red nodes : 25294 Black nodes : 74706 Per of red : 25.294
Trial Number : 7 Red nodes : 25248 Black nodes : 74752 Per of red : 25.247999999999998
Trial Number : 8 Red nodes : 25397 Black nodes : 74603 Per of red : 25.397
Trial Number : 9 Red nodes : 25582 Black nodes : 74418 Per of red : 25.582
Trial Number : 10 Red nodes : 25276 Black nodes : 74724 Per of red : 25.276
Trial Number : 11 Red nodes : 25321 Black nodes : 74679 Per of red : 25.320999999999998
Trial Number : 12 Red nodes : 25517 Black nodes : 74483 Per of red : 25.517
Trial Number : 13 Red nodes : 25440 Black nodes : 74560 Per of red : 25.44
Trial Number : 14 Red nodes : 25449 Black nodes : 74551 Per of red : 25.448999999999998
Trial Number : 15 Red nodes : 25446 Black nodes : 74554 Per of red : 25.446
Trial Number : 16 Red nodes : 25229 Black nodes : 74771 Per of red : 25.229000000000003
Trial Number : 17 Red nodes : 25288 Black nodes : 74712 Per of red : 25.288
Trial Number : 18 Red nodes : 25418 Black nodes : 74582 Per of red : 25.418000000000003
Trial Number : 19 Red nodes : 25338 Black nodes : 74662 Per of red : 25.338
Trial Number : 20 Red nodes : 25377 Black nodes : 74623 Per of red : 25.377
Trial Number : 21 Red nodes : 25456 Black nodes : 74544 Per of red : 25.456
Trial Number : 22 Red nodes : 25446 Black nodes : 74554 Per of red : 25.446
Trial Number : 23 Red nodes : 25371 Black nodes : 74629 Per of red : 25.371
Trial Number : 24 Red nodes : 25245 Black nodes : 74755 Per of red : 25.245
Trial Number : 25 Red nodes : 25324 Black nodes : 74676 Per of red : 25.324
Trial Number : 26 Red nodes : 25374 Black nodes : 74626 Per of red : 25.374000000000002
Trial Number : 27 Red nodes : 25306 Black nodes : 74694 Per of red : 25.306
Trial Number : 28 Red nodes : 25386 Black nodes : 74614 Per of red : 25.385999999999996
Trial Number : 29 Red nodes : 25339 Black nodes : 74661 Per of red : 25.339
Trial Number : 30 Red nodes : 25295 Black nodes : 74705 Per of red : 25.295
Trial Number : 31 Red nodes : 25462 Black nodes : 74538 Per of red : 25.462
Trial Number : 32 Red nodes : 25129 Black nodes : 74871 Per of red : 25.129
Trial Number : 33 Red nodes : 25465 Black nodes : 74535 Per of red : 25.465
Trial Number : 34 Red nodes : 25410 Black nodes : 74590 Per of red : 25.41
Trial Number : 35 Red nodes : 25214 Black nodes : 74786 Per of red : 25.214
Trial Number : 36 Red nodes : 25395 Black nodes : 74605 Per of red : 25.395
Trial Number : 37 Red nodes : 25378 Black nodes : 74622 Per of red : 25.378
Trial Number : 38 Red nodes : 25402 Black nodes : 74598 Per of red : 25.402
Trial Number : 39 Red nodes : 25243 Black nodes : 74757 Per of red : 25.243
Trial Number : 40 Red nodes : 25321 Black nodes : 74679 Per of red : 25.320999999999998
Trial Number : 41 Red nodes : 25402 Black nodes : 74598 Per of red : 25.402
Trial Number : 42 Red nodes : 25355 Black nodes : 74645 Per of red : 25.355
Trial Number : 43 Red nodes : 25511 Black nodes : 74489 Per of red : 25.511
Trial Number : 44 Red nodes : 25331 Black nodes : 74669 Per of red : 25.331
Trial Number : 45 Red nodes : 25325 Black nodes : 74675 Per of red : 25.324999999999996
Trial Number : 46 Red nodes : 25270 Black nodes : 74730 Per of red : 25.27
Trial Number : 47 Red nodes : 25313 Black nodes : 74687 Per of red : 25.313000000000002
Trial Number : 48 Red nodes : 25410 Black nodes : 74590 Per of red : 25.41
Trial Number : 49 Red nodes : 25386 Black nodes : 74614 Per of red : 25.385999999999996
Trial Number : 50 Red nodes : 25472 Black nodes : 74528 Per of red : 25.472
Trial Number : 51 Red nodes : 25219 Black nodes : 74781 Per of red : 25.219
Trial Number : 52 Red nodes : 25583 Black nodes : 74417 Per of red : 25.583
Trial Number : 53 Red nodes : 25451 Black nodes : 74549 Per of red : 25.451

Trial Number : 54 Red nodes : 25277 Black nodes : 74723 Per of red : 25.277
Trial Number : 55 Red nodes : 25576 Black nodes : 74424 Per of red : 25.576
Trial Number : 56 Red nodes : 25475 Black nodes : 74525 Per of red : 25.474999999999998
Trial Number : 57 Red nodes : 25453 Black nodes : 74547 Per of red : 25.453
Trial Number : 58 Red nodes : 25467 Black nodes : 74533 Per of red : 25.467000000000002
Trial Number : 59 Red nodes : 25595 Black nodes : 74405 Per of red : 25.595000000000002
Trial Number : 60 Red nodes : 25591 Black nodes : 74409 Per of red : 25.591
Trial Number : 61 Red nodes : 25261 Black nodes : 74739 Per of red : 25.261
Trial Number : 62 Red nodes : 25264 Black nodes : 74736 Per of red : 25.263999999999996
Trial Number : 63 Red nodes : 25205 Black nodes : 74795 Per of red : 25.205
Trial Number : 64 Red nodes : 25462 Black nodes : 74538 Per of red : 25.462
Trial Number : 65 Red nodes : 25401 Black nodes : 74599 Per of red : 25.401
Trial Number : 66 Red nodes : 25609 Black nodes : 74391 Per of red : 25.608999999999998
Trial Number : 67 Red nodes : 25403 Black nodes : 74597 Per of red : 25.403
Trial Number : 68 Red nodes : 25412 Black nodes : 74588 Per of red : 25.412000000000003
Trial Number : 69 Red nodes : 25420 Black nodes : 74580 Per of red : 25.419999999999998
Trial Number : 70 Red nodes : 25438 Black nodes : 74562 Per of red : 25.438
Trial Number : 71 Red nodes : 25409 Black nodes : 74591 Per of red : 25.409
Trial Number : 72 Red nodes : 25541 Black nodes : 74459 Per of red : 25.541000000000004
Trial Number : 73 Red nodes : 25309 Black nodes : 74691 Per of red : 25.308999999999997
Trial Number : 74 Red nodes : 25318 Black nodes : 74682 Per of red : 25.318
Trial Number : 75 Red nodes : 25452 Black nodes : 74548 Per of red : 25.452
Trial Number : 76 Red nodes : 25398 Black nodes : 74602 Per of red : 25.398
Trial Number : 77 Red nodes : 25375 Black nodes : 74625 Per of red : 25.374999999999996
Trial Number : 78 Red nodes : 25586 Black nodes : 74414 Per of red : 25.586
Trial Number : 79 Red nodes : 25363 Black nodes : 74637 Per of red : 25.363000000000003
Trial Number : 80 Red nodes : 25406 Black nodes : 74594 Per of red : 25.406000000000002
Trial Number : 81 Red nodes : 25294 Black nodes : 74706 Per of red : 25.294
Trial Number : 82 Red nodes : 25541 Black nodes : 74459 Per of red : 25.541000000000004
Trial Number : 83 Red nodes : 25399 Black nodes : 74601 Per of red : 25.399
Trial Number : 84 Red nodes : 25565 Black nodes : 74435 Per of red : 25.564999999999998
Trial Number : 85 Red nodes : 25376 Black nodes : 74624 Per of red : 25.375999999999998
Trial Number : 86 Red nodes : 25210 Black nodes : 74790 Per of red : 25.21
Trial Number : 87 Red nodes : 25321 Black nodes : 74679 Per of red : 25.320999999999998
Trial Number : 88 Red nodes : 25206 Black nodes : 74794 Per of red : 25.206
Trial Number : 89 Red nodes : 25532 Black nodes : 74468 Per of red : 25.532
Trial Number : 90 Red nodes : 25442 Black nodes : 74558 Per of red : 25.441999999999997
Trial Number : 91 Red nodes : 25309 Black nodes : 74691 Per of red : 25.308999999999997
Trial Number : 92 Red nodes : 25251 Black nodes : 74749 Per of red : 25.251
Trial Number : 93 Red nodes : 25461 Black nodes : 74539 Per of red : 25.461
Trial Number : 94 Red nodes : 25338 Black nodes : 74662 Per of red : 25.338
Trial Number : 95 Red nodes : 25506 Black nodes : 74494 Per of red : 25.506
Trial Number : 96 Red nodes : 25416 Black nodes : 74584 Per of red : 25.416
Trial Number : 97 Red nodes : 25545 Black nodes : 74455 Per of red : 25.545
Trial Number : 98 Red nodes : 25389 Black nodes : 74611 Per of red : 25.389
Trial Number : 99 Red nodes : 25326 Black nodes : 74674 Per of red : 25.325999999999997

For N = 1000000 random numbers , 100 trials result:

Trial Number : 0 Red nodes : 254008 Black nodes : 745992 Per of red : 25.4008
Trial Number : 1 Red nodes : 253198 Black nodes : 746802 Per of red : 25.319799999999997
Trial Number : 2 Red nodes : 253930 Black nodes : 746070 Per of red : 25.393
Trial Number : 3 Red nodes : 253679 Black nodes : 746321 Per of red : 25.3679
Trial Number : 4 Red nodes : 254218 Black nodes : 745782 Per of red : 25.4218

Trial Number : 5 Red nodes : 253783 Black nodes : 746217 Per of red : 25.3783
Trial Number : 6 Red nodes : 253380 Black nodes : 746620 Per of red : 25.338
Trial Number : 7 Red nodes : 253882 Black nodes : 746118 Per of red : 25.3882
Trial Number : 8 Red nodes : 253712 Black nodes : 746288 Per of red : 25.371199999999998
Trial Number : 9 Red nodes : 254039 Black nodes : 745961 Per of red : 25.4039
Trial Number : 10 Red nodes : 253802 Black nodes : 746198 Per of red : 25.380200000000002
Trial Number : 11 Red nodes : 254332 Black nodes : 745668 Per of red : 25.4332
Trial Number : 12 Red nodes : 253926 Black nodes : 746074 Per of red : 25.392599999999998
Trial Number : 13 Red nodes : 253339 Black nodes : 746661 Per of red : 25.3339
Trial Number : 14 Red nodes : 254010 Black nodes : 745990 Per of red : 25.401
Trial Number : 15 Red nodes : 253483 Black nodes : 746517 Per of red : 25.348300000000002
Trial Number : 16 Red nodes : 253740 Black nodes : 746260 Per of red : 25.374000000000002
Trial Number : 17 Red nodes : 254297 Black nodes : 745703 Per of red : 25.4297
Trial Number : 18 Red nodes : 253914 Black nodes : 746086 Per of red : 25.391399999999997
Trial Number : 19 Red nodes : 253553 Black nodes : 746447 Per of red : 25.355299999999996
Trial Number : 20 Red nodes : 253971 Black nodes : 746029 Per of red : 25.397100000000002
Trial Number : 21 Red nodes : 253732 Black nodes : 746268 Per of red : 25.3732
Trial Number : 22 Red nodes : 253781 Black nodes : 746219 Per of red : 25.378099999999996
Trial Number : 23 Red nodes : 254574 Black nodes : 745426 Per of red : 25.457400000000003
Trial Number : 24 Red nodes : 253137 Black nodes : 746863 Per of red : 25.3137
Trial Number : 25 Red nodes : 254103 Black nodes : 745897 Per of red : 25.410300000000003
Trial Number : 26 Red nodes : 254258 Black nodes : 745742 Per of red : 25.4258
Trial Number : 27 Red nodes : 253696 Black nodes : 746304 Per of red : 25.3696
Trial Number : 28 Red nodes : 253518 Black nodes : 746482 Per of red : 25.3518
Trial Number : 29 Red nodes : 253782 Black nodes : 746218 Per of red : 25.3782
Trial Number : 30 Red nodes : 253651 Black nodes : 746349 Per of red : 25.3651
Trial Number : 31 Red nodes : 253723 Black nodes : 746277 Per of red : 25.3723
Trial Number : 32 Red nodes : 254390 Black nodes : 745610 Per of red : 25.439
Trial Number : 33 Red nodes : 254276 Black nodes : 745724 Per of red : 25.4276
Trial Number : 34 Red nodes : 253895 Black nodes : 746105 Per of red : 25.389499999999998
Trial Number : 35 Red nodes : 254333 Black nodes : 745667 Per of red : 25.4333
Trial Number : 36 Red nodes : 254033 Black nodes : 745967 Per of red : 25.4033
Trial Number : 37 Red nodes : 253463 Black nodes : 746537 Per of red : 25.3463
Trial Number : 38 Red nodes : 254247 Black nodes : 745753 Per of red : 25.4247
Trial Number : 39 Red nodes : 253741 Black nodes : 746259 Per of red : 25.3741
Trial Number : 40 Red nodes : 253893 Black nodes : 746107 Per of red : 25.3893
Trial Number : 41 Red nodes : 253813 Black nodes : 746187 Per of red : 25.3813
Trial Number : 42 Red nodes : 254117 Black nodes : 745883 Per of red : 25.4117
Trial Number : 43 Red nodes : 254630 Black nodes : 745370 Per of red : 25.463
Trial Number : 44 Red nodes : 253659 Black nodes : 746341 Per of red : 25.365900000000003
Trial Number : 45 Red nodes : 253746 Black nodes : 746254 Per of red : 25.3746
Trial Number : 46 Red nodes : 254143 Black nodes : 745857 Per of red : 25.4143
Trial Number : 47 Red nodes : 253401 Black nodes : 746599 Per of red : 25.3401
Trial Number : 48 Red nodes : 253839 Black nodes : 746161 Per of red : 25.383899999999997
Trial Number : 49 Red nodes : 253701 Black nodes : 746299 Per of red : 25.3701
Trial Number : 50 Red nodes : 254028 Black nodes : 745972 Per of red : 25.4028
Trial Number : 51 Red nodes : 254036 Black nodes : 745964 Per of red : 25.403599999999997
Trial Number : 52 Red nodes : 253964 Black nodes : 746036 Per of red : 25.396400000000003
Trial Number : 53 Red nodes : 254078 Black nodes : 745922 Per of red : 25.4078
Trial Number : 54 Red nodes : 253538 Black nodes : 746462 Per of red : 25.3538
Trial Number : 55 Red nodes : 253357 Black nodes : 746643 Per of red : 25.3357
Trial Number : 56 Red nodes : 253572 Black nodes : 746428 Per of red : 25.357200000000002
Trial Number : 57 Red nodes : 253679 Black nodes : 746321 Per of red : 25.3679
Trial Number : 58 Red nodes : 254059 Black nodes : 745941 Per of red : 25.4059
Trial Number : 59 Red nodes : 253979 Black nodes : 746021 Per of red : 25.3979

Trial Number : 60 Red nodes : 253728 Black nodes : 746272 Per of red : 25.3728
Trial Number : 61 Red nodes : 253904 Black nodes : 746096 Per of red : 25.390400000000003
Trial Number : 62 Red nodes : 253714 Black nodes : 746286 Per of red : 25.3714
Trial Number : 63 Red nodes : 253609 Black nodes : 746391 Per of red : 25.360899999999997
Trial Number : 64 Red nodes : 253719 Black nodes : 746281 Per of red : 25.371899999999997
Trial Number : 65 Red nodes : 253819 Black nodes : 746181 Per of red : 25.3819
Trial Number : 66 Red nodes : 254076 Black nodes : 745924 Per of red : 25.407600000000002
Trial Number : 67 Red nodes : 254049 Black nodes : 745951 Per of red : 25.4049
Trial Number : 68 Red nodes : 253838 Black nodes : 746162 Per of red : 25.3838
Trial Number : 69 Red nodes : 253791 Black nodes : 746209 Per of red : 25.379099999999998
Trial Number : 70 Red nodes : 253696 Black nodes : 746304 Per of red : 25.3696
Trial Number : 71 Red nodes : 253668 Black nodes : 746332 Per of red : 25.3668
Trial Number : 72 Red nodes : 254154 Black nodes : 745846 Per of red : 25.415399999999998
Trial Number : 73 Red nodes : 254195 Black nodes : 745805 Per of red : 25.4195
Trial Number : 74 Red nodes : 254231 Black nodes : 745769 Per of red : 25.423099999999998
Trial Number : 75 Red nodes : 253814 Black nodes : 746186 Per of red : 25.3814
Trial Number : 76 Red nodes : 253996 Black nodes : 746004 Per of red : 25.3996
Trial Number : 77 Red nodes : 254201 Black nodes : 745799 Per of red : 25.4201
Trial Number : 78 Red nodes : 254098 Black nodes : 745902 Per of red : 25.4098
Trial Number : 79 Red nodes : 253891 Black nodes : 746109 Per of red : 25.3891
Trial Number : 80 Red nodes : 254517 Black nodes : 745483 Per of red : 25.4517
Trial Number : 81 Red nodes : 254012 Black nodes : 745988 Per of red : 25.401200000000003
Trial Number : 82 Red nodes : 253775 Black nodes : 746225 Per of red : 25.377499999999998
Trial Number : 83 Red nodes : 254118 Black nodes : 745882 Per of red : 25.4118
Trial Number : 84 Red nodes : 254022 Black nodes : 745978 Per of red : 25.402200000000004
Trial Number : 85 Red nodes : 253977 Black nodes : 746023 Per of red : 25.3977
Trial Number : 86 Red nodes : 253742 Black nodes : 746258 Per of red : 25.374200000000002
Trial Number : 87 Red nodes : 253789 Black nodes : 746211 Per of red : 25.378899999999998
Trial Number : 88 Red nodes : 254022 Black nodes : 745978 Per of red : 25.402200000000004
Trial Number : 89 Red nodes : 253889 Black nodes : 746111 Per of red : 25.388899999999996
Trial Number : 90 Red nodes : 253922 Black nodes : 746078 Per of red : 25.3922
Trial Number : 91 Red nodes : 253984 Black nodes : 746016 Per of red : 25.3984
Trial Number : 92 Red nodes : 254419 Black nodes : 745581 Per of red : 25.4419
Trial Number : 93 Red nodes : 253840 Black nodes : 746160 Per of red : 25.384
Trial Number : 94 Red nodes : 253698 Black nodes : 746302 Per of red : 25.369799999999998
Trial Number : 95 Red nodes : 254310 Black nodes : 745690 Per of red : 25.430999999999997
Trial Number : 96 Red nodes : 253673 Black nodes : 746327 Per of red : 25.367299999999997
Trial Number : 97 Red nodes : 253819 Black nodes : 746181 Per of red : 25.3819
Trial Number : 98 Red nodes : 253442 Black nodes : 746558 Per of red : 25.3442
Trial Number : 99 Red nodes : 254324 Black nodes : 745676 Per of red : 25.4324

From the above , it can be observed that the percentage of red nodes among total nodes is **around 25 Percent.**

Red Black Trees are binary search trees that are additionally constrained by 4 rules

1. each node is either red or black
2. the root is black
3. every red node must have either 0 or 2 black children
4. every root to null path must have the same number of black nodes

In our implementation we have used the left leaning Red- Black BST such that:

A. No node has two red links connected to it.

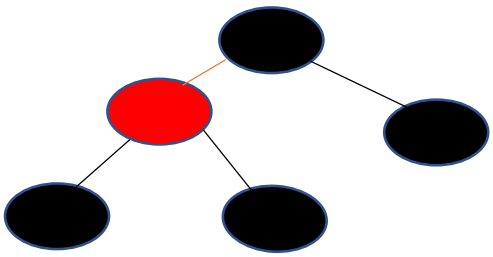
B. Every path from root to null link has the same number of black links.
C. Red links lean left.

Considering the above rules and properties:-

From the insertion it can be visualized that :-

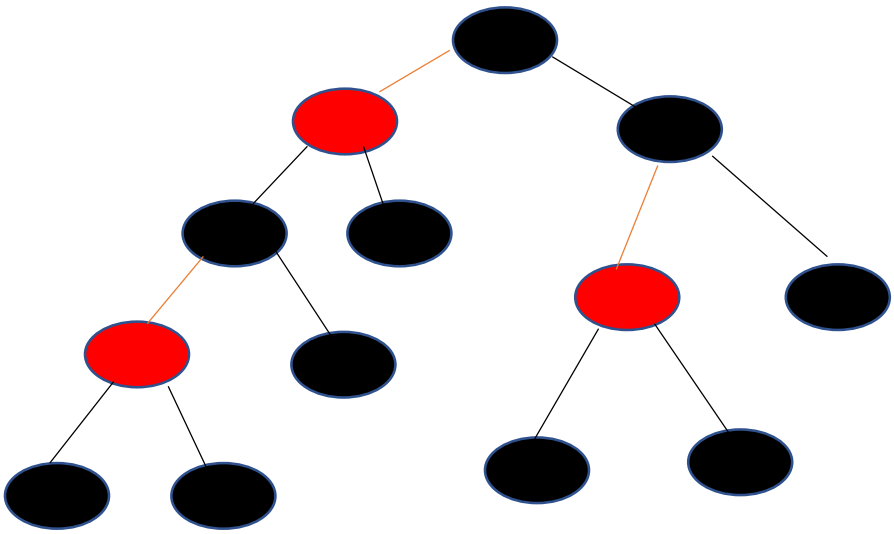
Smallest repetitive unit seems to be as shown below which satisfies all the condition of left leaning red black tree :-

A.



i.e. for every root black node one red node and three black nodes are incrementing and as this structure is almost replicating i.e. the 1 red node to 3 black nodes we can assume that red nodes : black nodes are in ratio 1:3 so red nodes are almost 25 percent of the total.

C. When this structure grows as it can be visualized that 3/13 = almost 25 percent(Note 1 is for the root black node , whose count i.e. 1 significance will become very less in large tree)



D. So as in our implementation as we are adding the red nodes as a new nodes in order to perform the flip , rotate operations in order to keep them balanced , we are also having red nodes around 25 percent

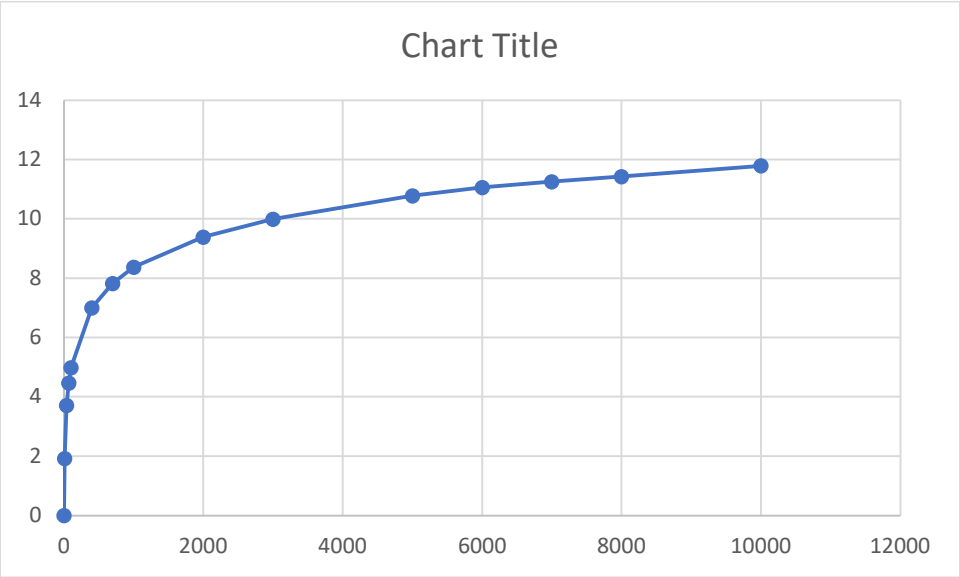
Q4. Run empirical studies to compute the average and std deviation of the average length of a path to a random node (internal path length divided by tree size) in a red-black BST built by insertion of N random keys into an initially empty tree, for N from 1 to 10,000. Do at least 1,000 trials for eachsize.

Solution:

Sample Data for various values of N with 1000 trials for each value of N is as below :-

S.No	N	Average(over 1000 trials) Length of Average Internal Path Length	Standard Deviation (over 1000 trials) Length of Average Internal Path Length
1.	1	0.0	0.0
2	10	1.9212000000000267	0.040872484632085246
3	40	3.708674999999992	0.10118235703421798
4	70	4.470814285714291	0.060762419575367814
5	100	4.982310000000005	0.09891392166929788
6	400	7.0040925000000005	0.06967932041682107
7	700	7.824718571428578	0.05625496215754604
8	1000	8.377903000000002	0.08153985277764493
9	2000	9.389327499999995	0.06407578710987487
10	3000	9.990862333333324	0.09757018808757346
11	5000	10.773706200000014	0.07398330001804461
12	6000	11.063357833333349	0.10404954180461348
13	7000	11.255885142857128	0.092539738577543
14	8000	11.430255125000023	0.05953402353368513
15	10000	11.790327200000027	0.06072219660190172

Graph of Average length Vs N



Output for the values of N is as shown below:

The average of the average path length of the tree of size : 1 is : 0.0
The std deviation of the average path length of the tree of size : 1 is : 0.0
The average of the average path length of the tree of size : 10 is : 1.9212000000000267
The std deviation of the average path length of the tree of size : 10 is : 0.040872484632085246
The average of the average path length of the tree of size : 40 is : 3.708674999999992
The std deviation of the average path length of the tree of size : 40 is : 0.10118235703421798
The average of the average path length of the tree of size : 70 is : 4.470814285714291
The std deviation of the average path length of the tree of size : 70 is : 0.060762419575367814
The average of the average path length of the tree of size : 100 is : 4.982310000000005
The std deviation of the average path length of the tree of size : 100 is : 0.09891392166929788
The average of the average path length of the tree of size : 400 is : 7.0040925000000005
The std deviation of the average path length of the tree of size : 400 is : 0.06967932041682107
The average of the average path length of the tree of size : 700 is : 7.806811428571424
The std deviation of the average path length of the tree of size : 700 is : 0.05625496215754604
The average of the average path length of the tree of size : 1000 is : 8.377903000000002
The std deviation of the average path length of the tree of size : 1000 is : 0.08153985277764493
The average of the average path length of the tree of size : 2000 is : 9.389327499999995
The std deviation of the average path length of the tree of size : 2000 is : 0.06407578710987487
The average of the average path length of the tree of size : 3000 is : 9.990862333333324
The std deviation of the average path length of the tree of size : 3000 is : 0.09757018808757346
The average of the average path length of the tree of size : 5000 is : 10.773706200000014
The std deviation of the average path length of the tree of size : 5000 is : 0.07398330001804461
The average of the average path length of the tree of size : 6000 is : 11.063357833333349
The std deviation of the average path length of the tree of size : 6000 is : 0.10404954180461348
The average of the average path length of the tree of size : 7000 is : 11.255885142857128
The std deviation of the average path length of the tree of size : 7000 is : 0.092539738577543
The average of the average path length of the tree of size : 8000 is : 11.430255125000023
The std deviation of the average path length of the tree of size : 8000 is : 0.05953402353368513
The average of the average path length of the tree of size : 10000 is : 11.790327200000027
The std deviation of the average path length of the tree of size : 10000 is : 0.06072219660190172

Note that the computation of values for all values of N i.e. 1 to 10000 is taking a lot of time. Even I have spent 24 hours and only the N up to 2000 results were computed.(O(N3) Algorithm 10000*1000* insertion of all N up to 10000 insertions)

I have kept the results of all the values of N which I have got in the separate file named q4_output.docx

Further for executing on the sample t_list please uncomment above for loop :

```
#for n in t_list:
```

And Comment the following :-

```
for m in range(10000):
        n=m+1
```

Q5. Implement the rank() and select() ordered operations for a BST. Use data set linked below. (i) What is the value of select(7) for the data set? (ii) What is the value of rank(7) for the data set?

Solution:

Output of the program is as shown :

Part a - Rank Operation
The result of the operation rank is : 6
Part b - Select Operation
The result of the operation select is : 8