Ripan Halder
NU ID: 001442993
PSA Assignment 4

Code: https://github.com/RipanHalder/INFO6205
Results: https://github.com/RipanHalder/INFO6205/tree/master/results/union_find
Submission PDF: https://github.com/RipanHalder/INFO6205 -> PSA Assignment 4 – Ripan Halder.pdf

Observations:

1. Quick Union Find for DEPTH vs SIZE
    a. DEPTH:

        When we are merging by depth/height, longer tree will always be a parent of a shorter tree. If the depth of two trees are same, then it is arbitrarily merged and only then the height is increased by 1.

        So, height of the tree will only increase when nodes count is doubled. So, we will get height which will be always less than or equal to log(N), where N is the number of nodes.

    b. SIZE:

        When we are merging by size, size of parent is increased with the size of child. Smaller node is attached to the larger node. So, the size of tree is always greater than or equal to 2^h where h is the height of the tree.

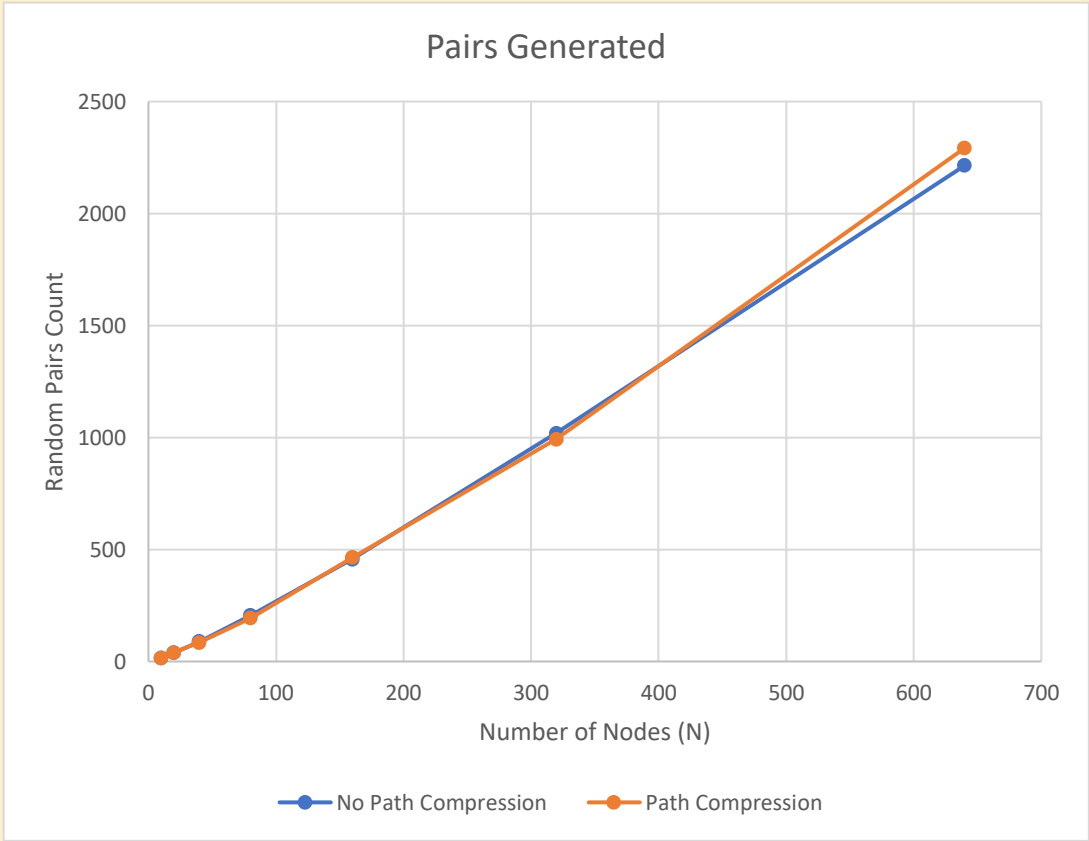
        So for DEPTH: $h <= \log_2(N)$

        And for SIZE: $N >= 2^h$ which implies to $\log_2(N) >= h$

        So, time complexity either by DEPTH or SIZE, will always be same. So we don't need to benchmark and test difference between them.

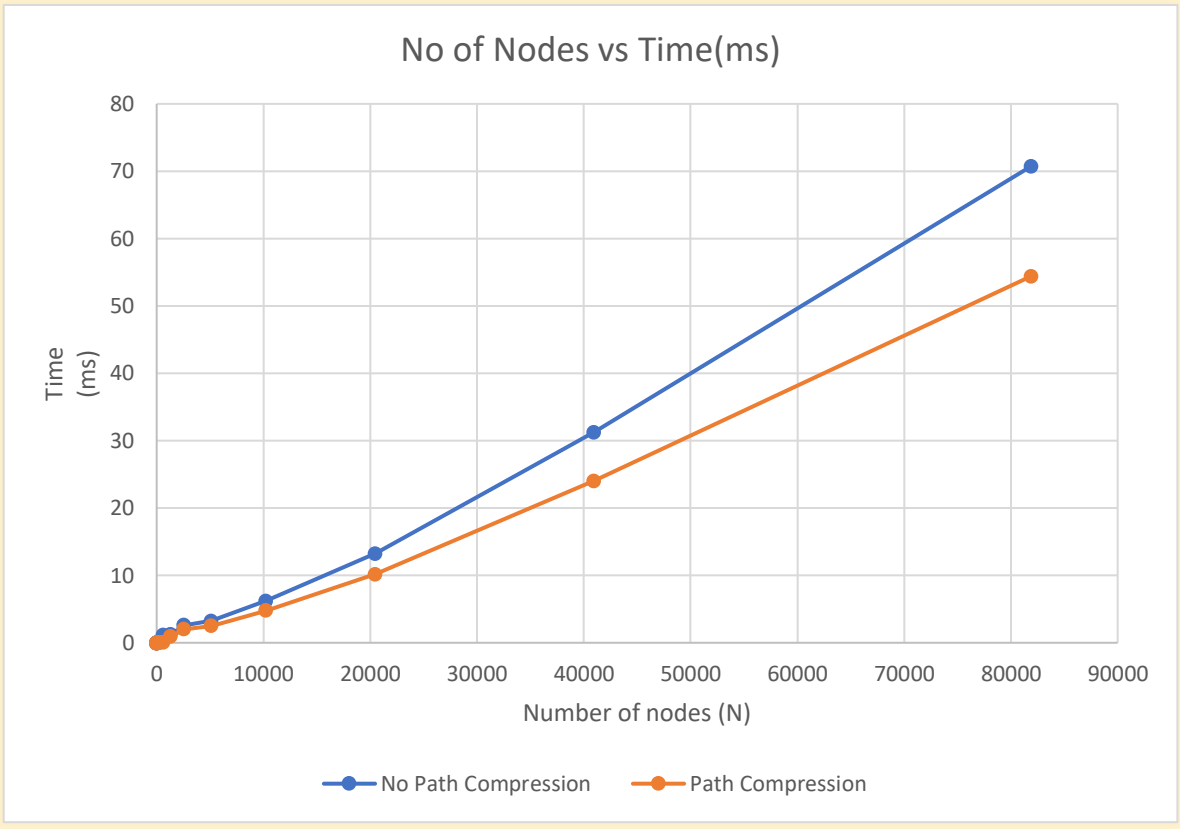
2. Comparing Path Compression vs No Compression on Pairs Generated:

N	Pairs by no path compression	Pairs by path compression
10	16.115	16.885
20	38.84	40.105
40	89.03	85.265
80	205.45	193.6
160	456.1	464.8
320	1019.595	992.23
640	2214.385	2292.59



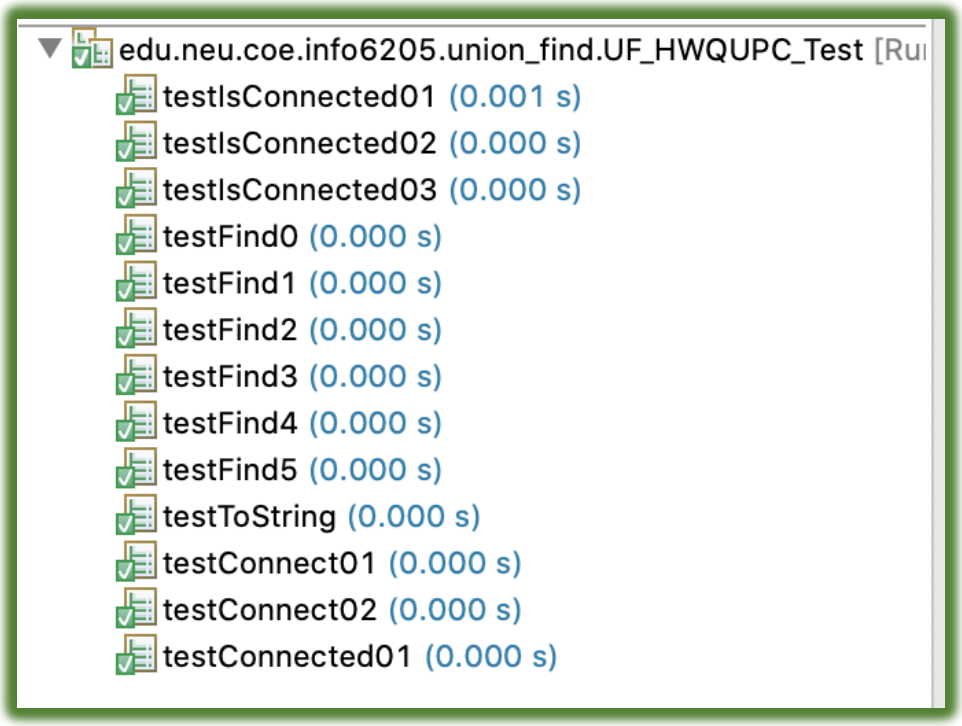
Benchmarking these results, we get the following Time (in ms) for random pairs generated through path compression or no path compression:

N	Time in ms (No Path Compression)	Time in ms (Path Compression)
10	0.000762832	0.01115321
20	0.001834399	0.002413665
40	0.005050571	0.00382086
80	0.018420168	0.008785985
160	0.067016609	0.019016515
320	0.278660571	0.04297637
640	1.128474538	0.09351898
1280	1.215267839	0.934821415
2560	2.623784728	2.018295945
5120	3.262748327	2.509806406
10240	6.21572638	4.781327985
20480	13.26378234	10.20290949
40960	31.26372345	24.04901804
81920	70.76361283	54.43354833

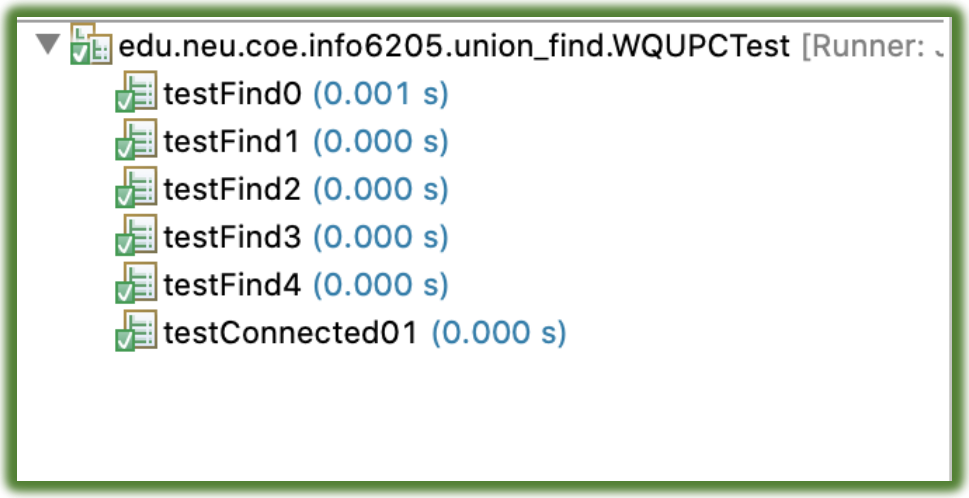


Unit Tests Screenshots:

1. UF_HWQUPC Test:



2. WQUPC Test:



Conclusion:

- *Random pairs generated during both union find using path compression or not is always almost similar.*

- *After, benchmarking we conclude that the algorithm works ~1.25 times faster when path compression is used versus when path compression is not used.*

- *Time complexity for is same for both depth versus size for this algorithm. Using any of them is fine but coding in size is bit easier rather than using approach of depth.*