Logan Kelsch

COSC 320 – Homework 07

4/17/24

Dr. Spickler

HOMEWORK 07 – REPORT

In this assignment, we gathered the code of AVL trees and Red-Black (RB) trees. With this code, we were tasked with implementing average IPL functions for both trees, and timing different methods of inserting and deleting nodes. More specifically, we timed how long it takes AVL trees and RB trees to insert a given number of nodes, as well as how long it takes to complete a series of cycles on the same tree, consisting of inserting and deleting a random node.

All collections consisted of 10 iterations of 3 variables, and gathered 8 data points for each iteration. These 8 data points were the average IPL and elapsed time for the insertions and cycles of both AVL and RB. I had four collection rounds, one with all 3 variables iterating, and three other collections with only two variables iterating. These three variables were the number of initial insertions (I), the number of insertion/deletion cycles (C) , and the maximum random value (rand_max). All iterations of all variables began at 50,000 and iterated by 50,000 (So all 10 iterations ran from 50,000 to 500,000).

The first collection of data was of all 3 variables iterating. Here is the table.

			AVL INSERT	AVL	AVL CYCLES	AVL	RB INSERT	RB	RB CYCLES	RB
RAND MAX	# OF NODES INITIAL	CYCLES	Avg IPL	Time	Avg IPL	Time	Avg IPL	Time	Avg IPL	Time
50000	50000	50000	15.386	0.03	15.0922	0.028	16.966	0.006	17.0154	0.04732
100000	100000	100000	16.3861	0.026	16.0612	0.065	18.0101	0.0143	18.0556	0.11019
150000	150000	150000	16.9628	0.042	16.6635	0.105	18.5778	0.0239	18.6079	0.18088
200000	200000	200000	17.4029	0.06	17.1121	0.15	19.0469	0.0355	19.0838	0.26299
250000	250000	250000	17.7286	0.081	17.4424	0.202	19.3705	0.0475	19.4236	0.35379
300000	300000	300000	17.9841	0.103	17.6839	0.252	19.6736	0.0604	19.7193	0.44441
350000	350000	350000	18.2361	0.128	17.9391	0.311	19.8538	0.0733	19.8965	0.54048
400000	400000	400000	18.3997	0.153	18.0945	0.368	20.0288	0.0889	20.067	0.64492
450000	450000	450000	18.65	0.178	18.3363	0.435	20.2332	0.1043	20.2736	0.76251
500000	500000	500000	18.744	0.204	18.4416	0.491	20.3941	0.1225	20.4457	0.86585

Two graphs were generated from this.

Average IPL vs. Number of Nodes, and Time (seconds) vs. Number of Nodes
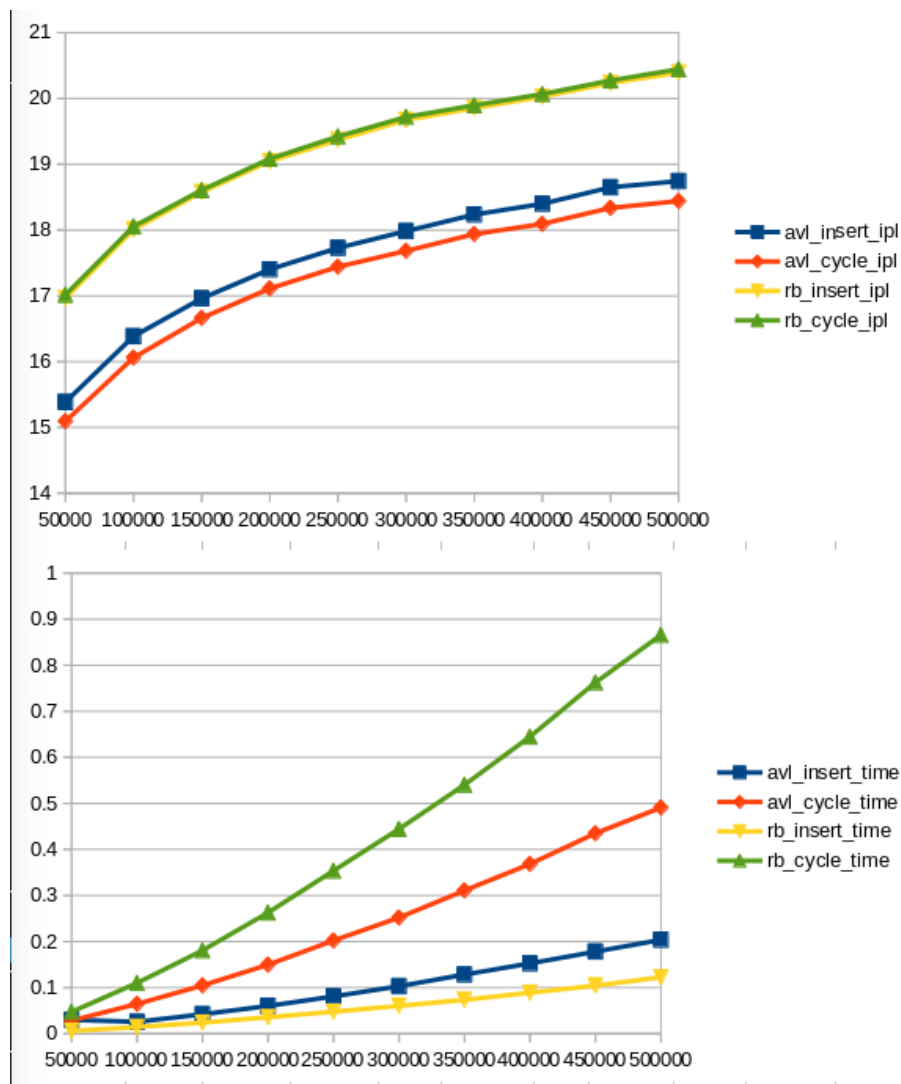
The second collection of data was of all but the initial insertion count (I) iterating. Here is the table.

	# OF NODES		AVL INSERT		AVL CYCLES		AVL		RB INSERT		RB CYCLES		RB
RAND MAX	INITIAL	CYCLES	Avg IPL	Time	Avg IPL	Time	Avg IPL	Time	Avg IPL	Time	Avg IPL	Time		
50000	200000	50000	16.0202	0.057	15.8091	0.023	19.0344	0.0316	19.0706	0.045				
100000	200000	100000	16.8506	0.057	16.5525	0.061	19.016	0.0348	19.0522	0.11				
150000	200000	150000	17.1874	0.062	16.8835	0.105	19.0233	0.0362	19.0693	0.182				
200000	200000	200000	17.4663	0.064	17.156	0.156	19.0356	0.0364	19.0868	0.267				
250000	200000	250000	17.5351	0.068	17.2472	0.213	19.0588	0.0365	19.1003	0.367				
300000	200000	300000	17.5774	0.068	17.3167	0.275	18.9866	0.0381	19.0301	0.475				
350000	200000	350000	17.6867	0.068	17.4263	0.342	19.0463	0.0373	19.0934	0.585				
400000	200000	400000	17.7182	0.069	17.4739	0.413	19.011	0.037	19.0541	0.709				
450000	200000	450000	17.7418	0.07	17.5243	0.487	19.0049	0.037	19.0678	0.849				
500000	200000	500000	17.7492	0.07	17.543	0.566	19.0521	0.037	19.1135	0.982				

Two graphs were generated from this.

Average IPL vs. Number of Nodes, and Time (seconds) vs. Number of Nodes
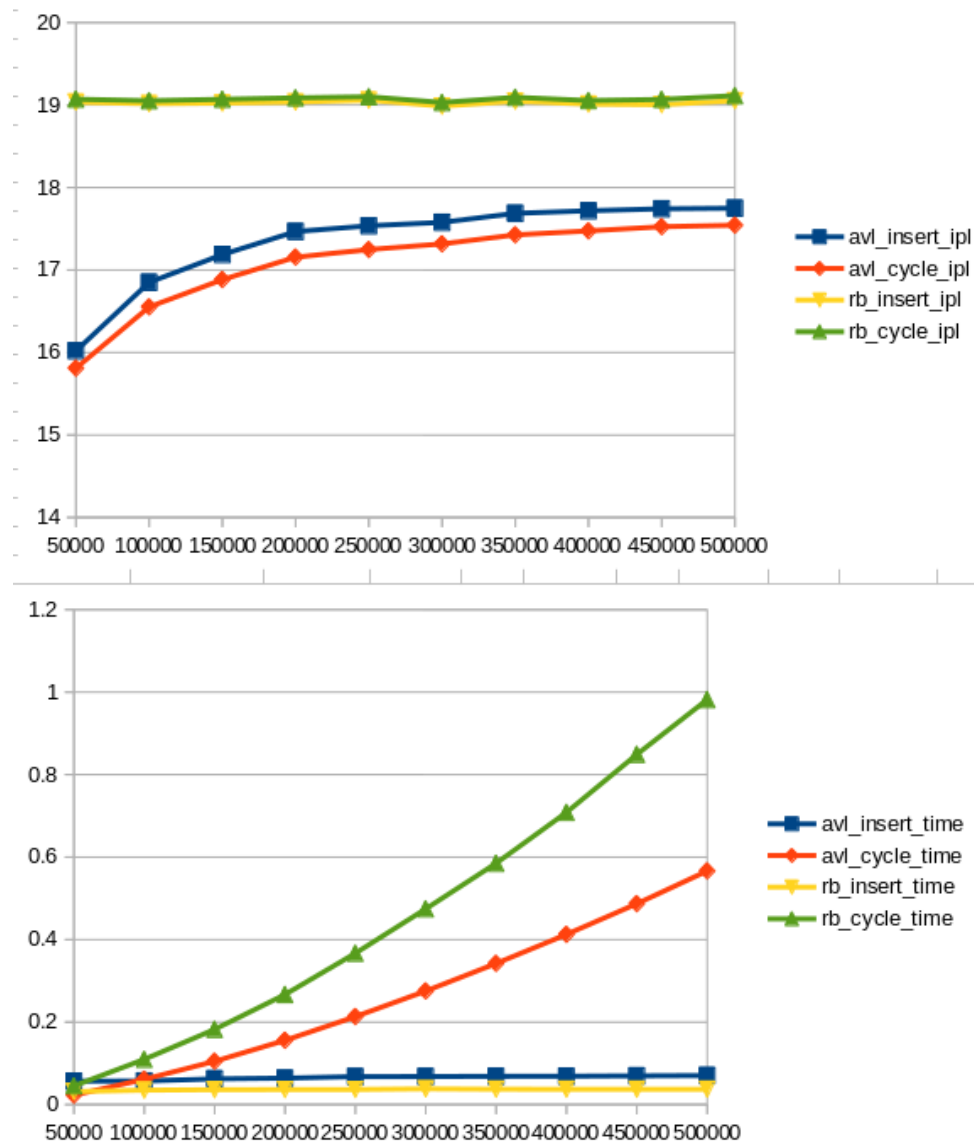
The third collection of data was of all but rand_max iterating. Here is the table.

			AVL INSERT		AVL CYCLES		RB INSERT		RB CYCLES	
RAND MAX	INITIAL	CYCLES	Avg IPL	Time	Avg IPL	Time	Avg IPL	Time	Avg IPL	Time
200000	50000	50000	15.8469	0.018	15.7303	0.046	16.9821	0.0061	17.0239	0.08098
200000	100000	100000	16.7029	0.027	16.5127	0.081	17.9762	0.0147	18.009	0.14078
200000	150000	150000	17.1063	0.043	16.8673	0.115	18.6242	0.0234	18.6719	0.19715
200000	200000	200000	17.388	0.06	17.0931	0.145	19.0966	0.0353	19.1455	0.25353
200000	250000	250000	17.5856	0.08	17.261	0.182	19.3352	0.048	19.3951	0.32199
200000	300000	300000	17.6909	0.1	17.3295	0.222	19.6258	0.0599	19.6793	0.38512
200000	350000	350000	17.757	0.121	17.3918	0.251	19.8436	0.0736	19.8977	0.44354
200000	400000	400000	17.839	0.142	17.4671	0.285	20.0328	0.0893	20.0956	0.50552
200000	450000	450000	17.9164	0.164	17.5426	0.315	20.2369	0.1071	20.3175	0.56149
200000	500000	500000	17.9284	0.183	17.5643	0.348	20.372	0.12	20.4624	0.6311

Two graphs were generated from this.

Average IPL vs. Number of Nodes, and Time (seconds) vs. Number of Nodes
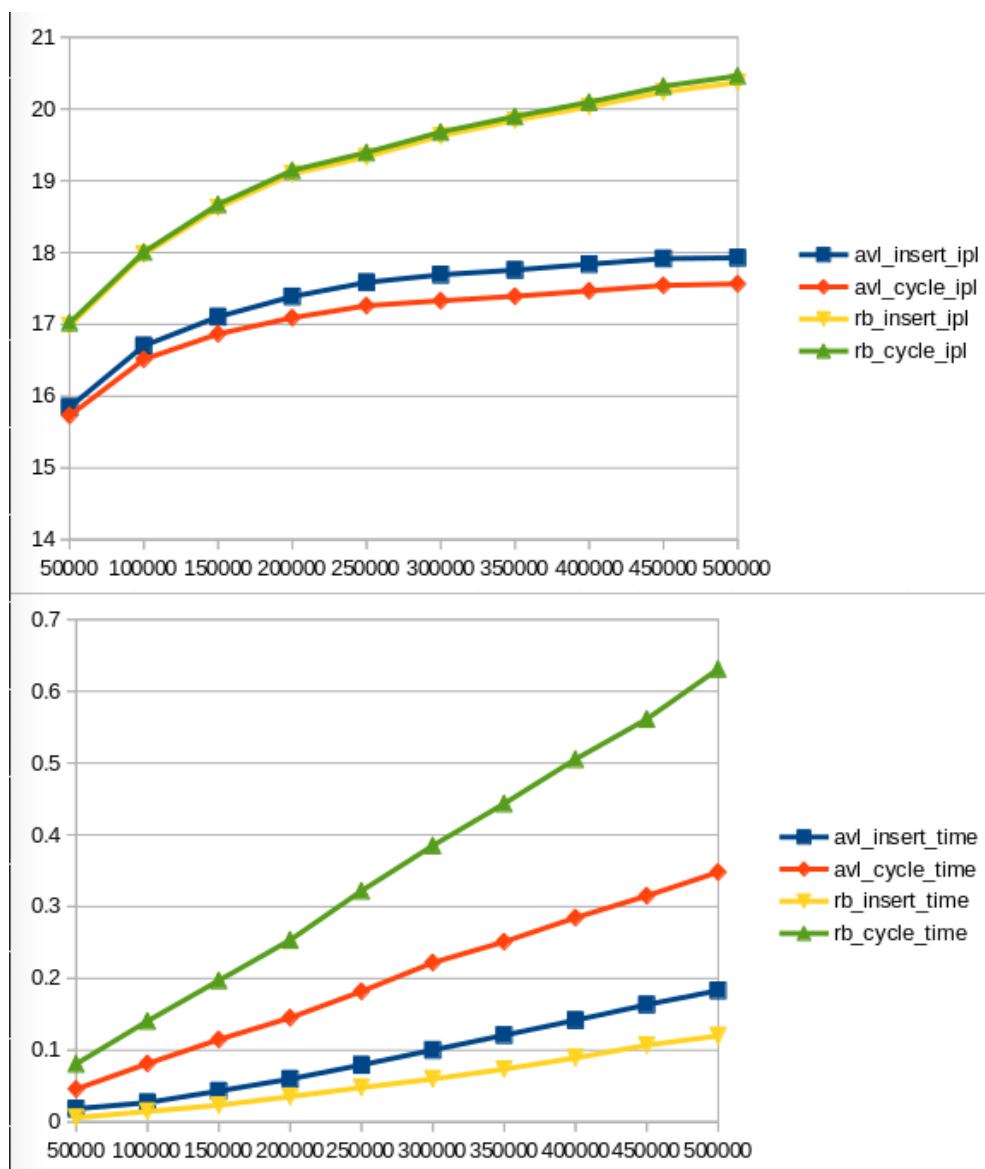
The fourth collection of data was of all but the cycle insertion count (C) iterating. Here is the table.

	# OF NODES		AVL INSERT		AVL CYCLES		RB INSERT		RB CYCLES	
RAND MAX	INITIAL	CYCLES	Avg IPL	Time	Avg IPL	Time	Avg IPL	Time	Avg IPL	Time
50000	50000	200000	15.3598	0.022	14.9794	0.11	16.9706	0.0096	17.033	0.184
100000	100000	200000	16.3576	0.026	16.0106	0.128	18.0036	0.0144	18.0638	0.215
150000	150000	200000	16.9844	0.042	16.655	0.138	18.5903	0.0239	18.6443	0.236
200000	200000	200000	17.3744	0.061	17.0747	0.149	19.003	0.036	19.0429	0.261
250000	250000	200000	17.7559	0.082	17.4865	0.161	19.341	0.0477	19.3846	0.279
300000	300000	200000	17.991	0.103	17.7403	0.171	19.6951	0.0614	19.7155	0.299
350000	350000	200000	18.2326	0.129	18.0106	0.18	19.8469	0.0786	19.8863	0.321
400000	400000	200000	18.4282	0.157	18.2134	0.188	20.0206	0.0891	20.0556	0.338
450000	450000	200000	18.6093	0.183	18.407	0.195	20.3202	0.107	20.3081	0.354
500000	500000	200000	18.7348	0.206	18.5475	0.197	20.3504	0.1243	20.3842	0.365

Two graphs were generated from this.

Average IPL vs. Number of Nodes, and Time (seconds) vs. Number of Nodes
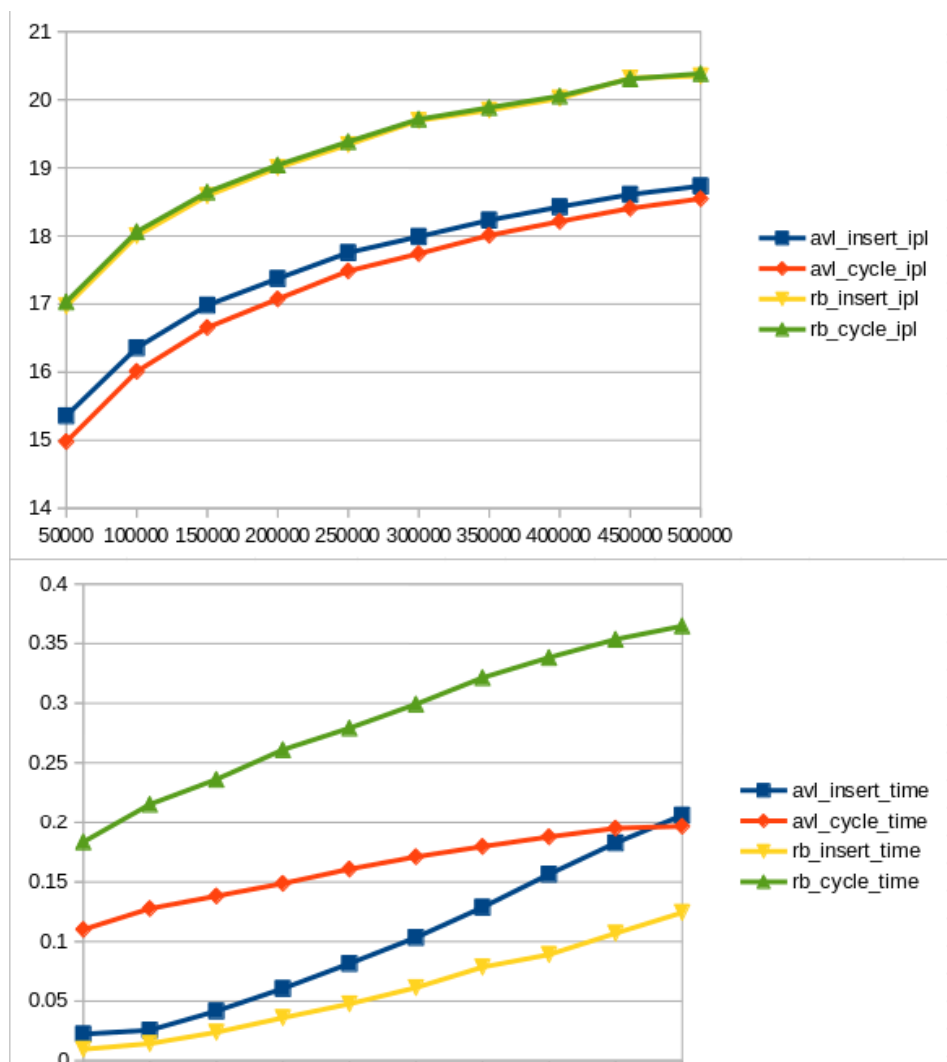
Given the excellent visualization we are able to generate, a few trends become much more noticeable when managing these variables. To begin, the tests consisting of all increasing values in the variables presents clear lg(n) plots, with AVL consistently having a lower average IPL value. In fact, the average IPL value after regular inserts and insert/delete cycles remains lower under AVL trees than RB trees under every circumstance, with any variable manipulation.

For each of the timing functions, the results remained exactly the same. With all time values increasing at an increasing rate, RB insert remains the fastest, followed by AVL insert, AVL cycles, and RB cycles every time.

This assignment clearly shows that, between AVL and RB trees, AVL trees are the most optimal, as the average IPL will predict the rate of all other tree-usage functions within a database. With a lower IPL every time, all other functions will run faster every time. The only condition in which RB trees would be more beneficial than AVL trees, is if the main use of the tree is insertion. Only under this condition, where speed of insertion is significantly more valuable than any other frequent function usage, should the RB tree be considered over the AVL tree.