

# Experimentation Assignment

Due: Wednesday 30th November 5pm

Email document to [pstuckey@unimelb.edu.au](mailto:pstuckey@unimelb.edu.au)

## A: Results Presentation

You have created three separate improvements for a method for solving problem P. The methods are A, B, and C. You have run experiments on a suite of 40 problem instances, 5 each from 8 different benchmark suites, for all combinations of baseline with the addition of 0,1,2 or 3 of the improvements.

The data for the experiments is shown overleaf, and given in

<http://people.eng.unimelb.edu.au/pstuckey/Experimentation.csv>

It shows the benchmark suite, the instance number, the size of the instance, and the times for all variations given in milliseconds.

Using just 1 A4 page present the experimental results, in the best possible manner. This should include discussion of the experimental results, in particular you should argue which of the methods A,B,C is the best improvement, and discuss any weakness of the results that you are presenting.

## B: Cross validation

In your own words, in about 3-4 paragraphs of text: define leave out  $p$  cross validation, and  $k$ -fold cross validation. Argue why cross validation is important, and which of the two methods is superior in which circumstances and why.

suite	bench	size	base	+A	+B	+C	+A+B+	+A+C	+B+C	+A+B+C
problem_10_10	1	10	8034	8220	7536	5236	8414	6665	5961	5231
problem_10_10	2	10	8923	8125	8324	7283	8726	7624	6507	4861
problem_10_10	3	10	12875	11954	10609	10628	11745	11237	9595	8577
problem_10_10	4	10	14023	12657	11412	9623	12688	11166	9521	7666
problem_10_10	5	10	15883	14608	12927	10814	14691	12708	9930	8712
Shaw_20_20	6	20	17970	16673	15237	11502	16729	14113	11602	11710
Shaw_20_20	7	20	23954	21877	19472	17412	21601	19640	16247	15261
Shaw_20_20	8	20	28324	25671	24185	20447	26257	23010	18963	16967
Shaw_20_20	9	20	36386	33465	30891	24667	34056	29018	24603	20950
Shaw_20_20	10	20	23330	21220	19917	15696	21440	18392	15496	13230
wbo_10_20	11	10	5921	6038	5130	5326	6079	5757	4091	3574
wbo_10_20	12	10	7352	6749	7090	6000	7291	6341	5917	5049
wbo_10_20	13	10	4324	4330	4361	2417	4763	3324	3355	2607
wbo_10_20	14	10	2323	2703	1617	1483	2248	2144	1660	1093
wbo_10_20	15	10	5634	5303	4513	5122	5389	5161	3744	2576
wbp_15_30	16	15	9354	8491	8106	5688	8722	6996	6241	6090
wbp_15_30	17	15	10203	9523	8439	8910	9469	9118	7950	6266
wbp_15_30	18	15	13212	12404	12215	7781	13149	10168	8866	8079
wbp_15_30	19	15	8939	8398	7651	4476	8369	6414	5528	4898
wbp_15_30	20	15	7823	7361	7523	5360	7583	6334	5595	3780
wbop_30_30	21	30	22562	20910	19009	15139	21229	17971	14765	13257
wbop_30_30	22	30	23132	21750	20283	15365	22130	18639	15840	14112
wbop_30_30	23	3	23423	21578	19573	17987	21521	19823	16538	14276
wbop_30_30	24	30	23231	21301	20522	15874	21919	18569	15572	14645
wbop_30_30	25	30	23434	21616	19987	15288	21816	18507	15416	13558
problem_40_20	26	40	28888	26059	24621	18320	26437	22156	18592	17545
problem_40_20	27	40	34234	31732	28683	22825	32030	27375	22011	20888
problem_40_20	28	40	34523	31938	29587	26035	32471	28962	23689	20116
problem_40_20	29	40	38455	35262	31750	27501	35491	31401	25725	22735
problem_40_20	30	40	39823	35967	34172	27349	36988	31559	26906	24124
wbp_30_15	31	30	22985	21546	18929	17245	21313	19437	15755	13455
wbp_30_15	32	30	20122	18387	16499	13396	18373	15944	13471	12470
wbp_30_15	33	30	20231	18716	18039	13937	19447	16338	13474	12727
wbp_30_15	34	30	19873	18599	17750	13245	19293	15954	13899	11317
wbp_30_15	35	30	25001	23173	20331	17034	22977	20144	16307	15401
wbo_30_15	36	30	23201	21361	19109	14430	21479	17862	14238	13844
wbo_30_15	37	30	19002	17530	16862	11929	18332	14720	12621	10822
wbo_30_15	38	30	18003	17117	15048	11366	16887	14168	11708	10618
wbo_30_15	39	30	18367	16772	15285	13225	16702	15055	12108	10745
wbo_30_15	40	30	18642	17650	16485	14781	18142	16216	13831	12099