Decision procedures and verification - DPLL

Jaroslav Šafář

August 24, 2020

1 Experimental evaluation

For experimental evaluation some smaller examples from SATLIB Benchmark Problems were used.

1.1 SAT examples

In the case of SAT problems of sizes 20, 50, 75, and 100, the first ten examples were selected for the experimental evaluation. Their resulting CPU times were averaged, and the results can be found in the following figure [1].

size	time-dpll	time-dpll2	time-dpll3	SAT
20	0.037170	0.001062	0.003131	True
50	0.332587	0.009645	0.046383	True
75	2.911390	0.112283	0.755583	True
100	17.291338	0.515635	4.602846	True

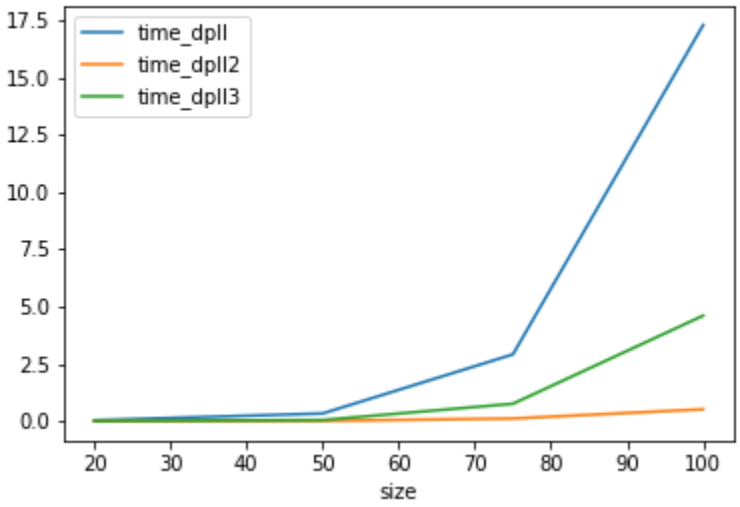

Figure 1: Averadge CPU time of SAT examples

1.2 UNSAT examples

In the case of UNSAT problems of sizes 50, 75, and 100, the first ten examples were selected for the experimental evaluation. Their resulting CPU times were again averaged, and the results can be found in the following figure [2].

size	time-dpll	time-dpll2	time-dpll3	SAT
50	0.871457	0.038526	0.161259	False
75	8.645794	0.313779	1.961455	False
100	52.235100	1.731449	14.286525	False

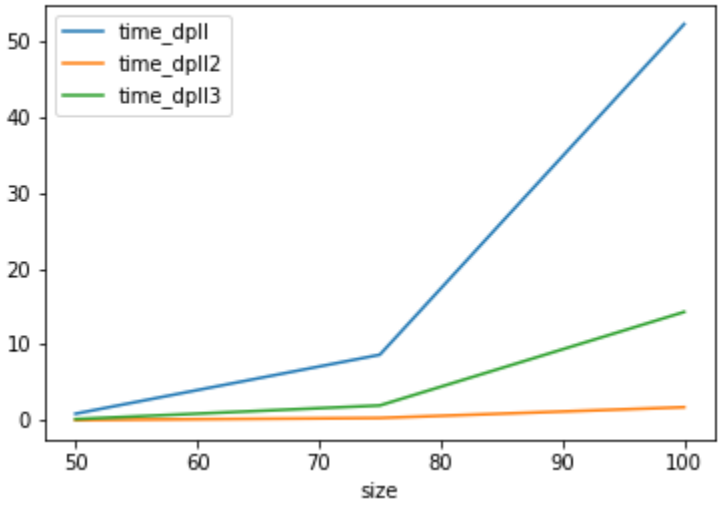

Figure 2: Averadge CPU time of UNSAT examples

1.3 Conclusion

As expected, the relationship of CPU time on the size of the formula looks exponential. That is because the number of possible assignments for the formula with n variables is 2^n. In the worst case, the DPLL algorithm would need to examine all of them to determine whether the formula is satisfiable or not. Also, the SAT problem is known to be (first) NP-complete problem (Cook-Levin theorem).